Research Paper

Volume : 2 | Issue : 8 | August 2013 • ISSN No 2277 - 8179

Achieving Network Level Privacy in Wireless Sensor Networks

Engineering

KEYWORDS : : anonymity; eavesdropping; hop-by-hop trace back; privacy; routing; wireless sensor networks.

THATIPAMULA RAJU — M. TECH COMPUTER SCIENCE, 11X31D0507(Roll No), RVR INSTITUTE OF ENGG.& TECH., SHERIGUDA, IBRAHIMPATNAM, RANGAREDDY, ANDHRAPRADESH

D. DEEPIKA RANI — RVR INSTITUTE OF ENGG. & TECH., SHERIGUDA, IBRAHIMPATNAM, RANGAREDDY, ANDHRAPRADESH

ABSTRACT

Full network level privacy has often been categorized into four sub-categories: Identity, Route, Location and Data privacy. Achieving full network level privacy is a critical and challenging problem due to the constraints imposed by the sensor nodes (e.g., energy, memory and computation power), sensor networks (e.g., mobility and topology) and QoS issues (e.g., packet reach-ability and timeliness). In this paper, we proposed two new identity, route and location privacy algorithms and data privacy mechanism that addresses this problem. The proposed solutions provide additional trustworthiness and reliability at modest cost of memory and energy. Also, we proved that our proposed solutions provide protection against various privacy disclosure attacks, such as eavesdropping and hop-by-hop trace back attacks.

1.Introduction

With the spreading application of Wireless Sensor Networks (WSNs) in various sensitive areas such as health-care, military, habitat monitoring, *etc*, the need to ensure security and privacy is becoming imperatively important. For example, in battlefield application scenario, "the location of a soldier should not be exposed if he initiates broadcast query" . In the meantime, query must be transferred to the destination in an encrypted manner via only trusted en-route nodes. Similarly, in habitat monitoring application scenarios, such as Great Duck Island or Save-the-panda application where large numbers of sensor nodes are deployed to observe the vast habitat of ducks and pandas, an adversary can try to capture the panda or duck by back-tracing the routing path until it reaches the source sensor nodes. Therefore, in order to prevent the adversary from back-tracing, the route, location and data privacy mechanisms must be enforced. With respect to these application scenarios, network level privacy has often been categorized into four categories:

1. Sender node identity privacy: no intermediate node can get any information about who is sending the packets except the source, its immediate neighbors and the destination,
2. Sender node location privacy: no intermediate node can have any information about the location (in terms of physical distance or number of hops) about the sender node except the source, its immediate neighbors and the destination,
3. Route privacy: no node can predict the information about the complete path (from source to destination). Also, a mobile adversary gets no clue to trace back the source node either from the contents and/or directional information of the captured packet(s), and
4. Data packet privacy: no node can see the information inside in a payload of the data packet except the source and the destination.

In order to achieve this goal, we incorporate basic design features from related research fields such as geographic routing and cryptographic systems. To our knowledge, we propose the first full network level privacy solution for WSNs. Our contribution lies in following features.

- A new Identity, Route and Location (IRL) privacy algorithm is proposed that ensures the anonymity of source node's identity and location. It also assures that the packets will reach their destination by passing through only trusted intermediate nodes.
- A new reliable Identity, Route and Location (r-IRL) privacy algorithm is proposed, which is the extension of our proposed IRL algorithm. This algorithm has the ability to forward packets from multiple secure paths to increase the packet reach-ability.
- A new data privacy mechanism is proposed, which is unique in the sense that it provides data secrecy and packet authentication *in the presence of identity anonymity*.

2. Network, Assumptions and Adversary Model

2.1. Network Model

A wireless sensor network (WSN) is composed of large number of small sensor nodes that are of limited resource and densely deployed in an environment. Whenever end users require information about any event related to some object(s), they send a query to the sensor network via the base station. And the base station propagates that query to the entire network or to a specific region of the network.

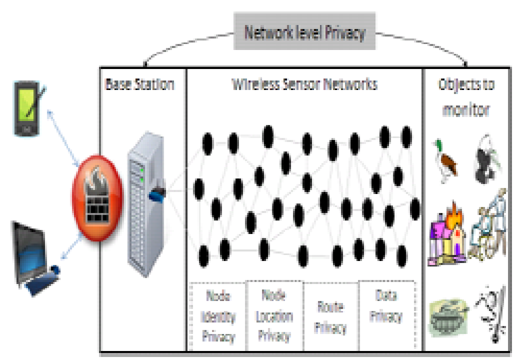


Figure 1. Typical WSN scenario.

2.2. Adversary Model

- Device-rich: the adversary is equipped with devices like antenna and spectrum analyzers, so that the adversary can measure the angle of arrival of the packet and received signal strength
- Resource-rich: the adversary has no resource constraint in computation power, memory or energy.

3. Proposed Scheme

3.1. Concepts and Definitions

In our proposed algorithms, we have used two notions: direction and trust. Both these notions (direction and trust) are used to provide reliable (non-malicious and non-faulty) secure paths for achieving robust route privacy. **Direction**: The first notion used in our algorithms is that of direction. The physical location of the base station is the reference point for each sensor node. Based on this reference point, each node classifies its neighboring nodes into four categories: (1) forward neighboring nodes (*F*), (2) right side backward neighboring nodes (*Br*), (3) left side backward neighboring nodes (*Bl*), and (4) middle backward neighboring nodes (*Bm*).

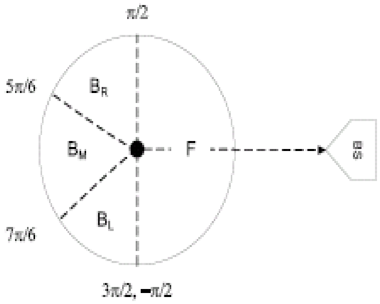


Figure 2. Neighbor node classification

Trust: The second notion used in our algorithms is that of trust. The definition of a trust here is based on our other paper and restated here. A node can be classified into one of the three categories trustworthy, untrustworthy, and uncertain.

Based on these successful and unsuccessful interactions node x can calculate the trust value of node y in following fashion:

$$T_{x,y} = \left[100\left(\frac{S_{x,y}}{S_{x,y}+U_{x,y}}\right)\left(1-\frac{1}{S_{x,y}+1}\right)\right] \qquad (2)$$

$$Mp(T_{x,y}) = \left\{\begin{array}{ll} \text{trustworthy} & 100-f \leq T_{x,y} \leq 100 \\ \text{uncertain} & 50-g \leq T_{x,y} < 100-f \\ \text{untrustworthy} & 0 \leq T_{x,y} < 50-g \end{array}\right\}.$$

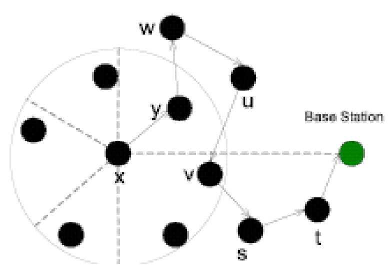
Algorithm 1 IRL - Routing at Source Node.

1: $prev_{hop} \leftarrow \emptyset$; $next_{hop} \leftarrow \emptyset$;
2: **if** $M(t_F) \neq \emptyset$ **then**
3: $next_{hop}(k)$ = Rand($M(t_F)$);
4: **else**
5: **if** $M(t_{B_r}) \cup M(t_{B_l}) \neq \emptyset$ **then**
6: $next_{hop}(k)$ = Rand($M(t_{B_r}) \cup M(t_{B_l})$);
7: **else if** $M(t_{B_m}) \neq \emptyset$ **then**
8: $next_{hop}(k)$ = Rand($M(t_{B_m})$);
9: **else**
10: Drop packet and Exit;
11: **end if**
12: **end if**
13: Set $prev_{hop} = my_{id}$;
14: Form pkt $p = \{prev_{hop}, next_{hop}, seqID, payload\}$;
15: Create Signature and save in buffer;
16: Forward packet to $next_{hop}$;
17: Set timer $\Delta t = \frac{D}{d_{next_{hop}}} \times p_t$;
18: **while** $\Delta t = true$ **do**
19: Signature remains in buffer;
20: **end while**
21: Signature removed from buffer;

Algorithm 2 IRL - Routing at Intermediate Node.

1: $next_{hop} \leftarrow \emptyset$;
2: $M_{temp} = \emptyset$
3: **if** Signature of new packet already exists in buffer **then**
4: $M_{temp} = \{M_{temp}\} + LasttimePrev_{hop}$
5: $M_{temp} = \{M_{temp}\} + LasttimeNext_{hop}$
6: Set $counter = timesReceviedBefore + 1$;
7: Remove signature from buffer;
8: **if** $counter = 3$ **then**
9: Drop packet and exit;
10: **end if**
11: **end if**
12: $M_{temp} = \{M_{temp}\} + prev_{hop}$
13: **if** $(M(t_F) - \{M(t_F) \bigcap M_{temp}\}) \neq \emptyset$ **then**
14: $next_{hop}(k)$ = Rand($M(t_F) - \{M(t_F) \bigcap M_{temp}\}$);
15: **else**
16: **if** packet came from B_r **then**
17: $M_{temp1} = M(t_{B_l}) \cup M(t_{B_m})$
18: **if** $M_{temp1} \neq \emptyset$ **then**
19: $next_{hop}(k)$ = Rand(M_{temp1});
20: **else if** $M(t_{B_r}) \neq \emptyset$ **then**
21: $next_{hop}(k)$ = Rand($M(t_{B_r}) - \{M(t_{B_r}) \bigcap M_{temp}\}$);
22: **else**
23: Drop packet and Exit;
24: **end if**
25: **else if** packet came from B_l **then**
26: $M_{temp2} = M(t_{B_r}) \cup M(t_{B_m})$
27: **if** $M_{temp2} \neq \emptyset$ **then**
28: $next_{hop}(k)$ = Rand($M_{temp2} - \{M_{temp2} \bigcap M_{temp}\}$);
29: **else if** $M(t_{B_l}) \neq \emptyset$ **then**
30: $next_{hop}(k)$ = Rand($M(t_{B_l}) - \{M(t_{B_l}) \bigcap M_{temp}\}$);
31: **else**
32: Drop packet and Exit;
33: **end if**
34: **else**
35: $M_{temp3} = M(t_{B_r}) \cup M(t_{B_l})$
36: **if** $M_{temp3} \neq \emptyset$ **then**
37: $next_{hop}(k)$ = Rand($M_{temp3} - \{M_{temp3} \bigcap M_{temp}\}$);
38: **else if** $M(t_{B_m}) \neq \emptyset$ **then**
39: $next_{hop}(k)$ = Rand($M(t_{B_m}) - \{M(t_{B_m}) \bigcap M_{temp}\}$);
40: **else**
41: Drop packet and Exit;
42: **end if**
43: **end if**
44: **end if**
45: Rest is same as Algorithm 1 from lines 13:21;

Figure 3. Sample routing scenario of IRL scheme.



This routing strategy may result in the creation of a cycle (loop). However, due to the randomness in the selection of the next-hop and the presence of the different four direction sets, the probability of creation of any cycle is very low. Nevertheless, in order to fully avoid the occurrence of the cycles, each node (prior to forwarding of a packet) will save the signature of the packet in the buffer for the δt time, that is:

$$\delta t = 2\left(\frac{D}{d} \times p_t\right)$$

where *D* is the distance between the forwarding node and the base station, *d* is the distance between the forwarding node and the next hop, and *pt* is the propagation transfer time between the forwarding node and the next hop. This signature consists of two fields: (1) sequence number of the packet, and (2) the pay-

Volume : 2 | Issue : 8 | August 2013 • ISSN No 2277 - 8179

load. Corresponding to this signature, three more fields are also stored in the buffer: (1) previous hop identity, (2) next hop identity where the packet is forwarded, and (3) counter, that tells how many times the same packet is received by the node.

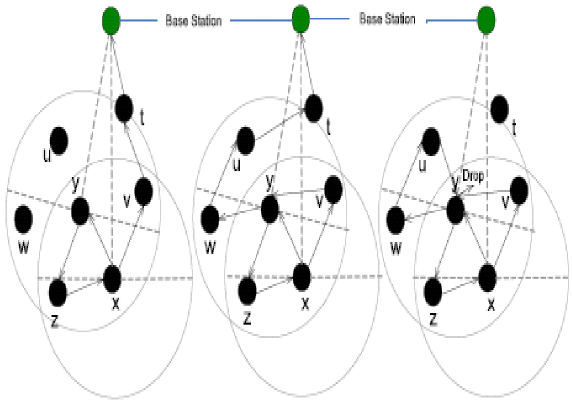


Figure 4. Three sample cycle detection and prevention scenarios.

Our proposed data privacy approach provides several benefits. Firstly, data secrecy is achieved in the presence of identity anonymity. This feature is not available in earlier proposed privacy schemes. Secondly, the base station will receive both the identity of the actual source node and message authentication. If the packet has been successfully decrypted with the shared secret key, it means that packet is received from genuine sensor node.

4. Analysis and Evaluation
4.1. Security Resiliency Analysis

Suppose we have an adversary *A* who strives to defeat our privacy protocols and guess the original source node. We will distinguish between two kinds of nodes. A source node is the original sender of a packet *q* and a forwarding node is the node that forwards a packet to another node until it reaches the destination. Hence the source node is also a forwarding node.

We will deal with separate cases. Given a packet *q* and a subset of nodes *N′*, find out the sender node *s*. In other words, the algorithm for the adversary takes two inputs and outputs a node *s′*; Namely *A*(*q*,*N′*) = *s′*. If *s′* = *s*, the adversary succeeds in defeating our protocol. We have to find:

Pr[*A*(*q*,*N′*) = *s*], which is the probability for an adversary to find out the sender node. Our assumption I s that, from an adversarial perspective, all nodes are equally likely to be senders of a packet. This does not necessarily mean that the network traffic is uniformly distributed. Notice that if the adversary knows beforehand which nodes are more likely to send packets, then no privacy preserving method can

adversary will know. This is true since the adversary can see all incoming packets to the node m and to its neighbor nodes (the forward and the backward nodes). Thus it can see if the payload of q is not equal to the payload of any q' being received by these nodes in a given interval of time. If this is the case, then the adversary will know the sender.

Now if none of the nodes in N_m are the senders, then the packet was forwarded by a node i that is two hops away from m. The adversary knows the ID of that node through the packet q. Thus the adversary makes a list of all the possible backward nodes in the backward set of i. Let that number be denoted by c. Notice that node i could also be the possible sender. Hence the total number of possible senders would be $c+1$. We have:

$$\Pr[A(q,N)=s] = \Pr[A(q,N)=s \mid s\in N_m]\Pr[s\in N_m] +$$

$$< \frac{m_f+m_b+1}{N} + \frac{1}{c+1}\left(1-\frac{m_f+m_b+1}{N}\right)$$

∎

Now, suppose the adversary is in possession of two nodes at the same time; m_1 and m_2. We can safely assume that $N_{m_1}\cap N_{m_2}=\varphi$, since it would be more advantageous to the adversary to cover nodes with non-overlapping radio ranges. The adversary will always know whenever any node in N_{m_1} or N_{m_2} is the sender of a packet. How about the case when they are not the senders? There could be two possible cases: without loss of generality, first assume that $m_2 \in C_{m_1}$. If the packet q was received by some node in N_{m_1} and was received by some node in N_{m_2} before, then the adversary had already checked it when the packet was sent to a node in N_{m_2}. Thus the adversary need only check packets received in N_{m_1} that were not received by N_{m_2}. In this case, the sender cannot be in N_{m_2}. In any case, the adversary has to find out the backward sets of $\rightarrow^2 m_1$ or $\rightarrow^2 m_2$, depending on where the packet was received. Since, in the adversary's knowledge, all nodes are equally likely to be senders, the probability of a packet being received at the two sets is the same. In case $m_2 \notin C_{m_1}$, then the adversary has no real advantage except that it can see packets at two disjoint locations in the network. Thus we only state the case when $m_2 \in C_{m_1}$. We have the following result:

Claim 3: Suppose the adversary is in possession of two nodes m_1 and m_2. Assume further that $m_2 \in C_{m_1}$. Let $c_1 = |C_{\rightarrow^2 m_1}|$ and $c_2 = |C_{\rightarrow^2 m_2}|$ then:

$$\Pr[A(q,N)=s] = \frac{|N_{m_1}|+|N_{m_2}|}{N} + \frac{1}{2}\left(\frac{1}{c_1+1-|N_{m_2}|}+\frac{1}{c_2+1}\right)\left(1-\frac{|N_{m_1}|+|N_{m_2}|}{N}\right) \quad (9)$$

In general, we have:

Claim 4: Let us assume that A is in possession of k nodes $m_k \rightarrow^{r_1} \cdots \rightarrow^{r_{k-2}} m_2 \rightarrow^{r_{k-1}} m_1$ and let m_f and m_b denote the average number of forward and backward nodes averaged over all the k nodes. Let $t = m_f + m_b + 1$. Let for $1 \le i \le k$, $c_i = |C_{\rightarrow^2 m_i}|$, then:

$$\Pr[A(q,N)=s] = \frac{kt}{N} + \frac{1}{k}\left(\frac{1}{c_1+1-(k-1)t}+\frac{1}{c_2+1-(k-2)t}+\cdots+\frac{1}{c_k+1}\right)\left(1-\frac{kt}{N}\right) \quad (10)$$

4.2. Memory Consumption Analysis

Each sensor node needs to maintain one table that contains the list of neighboring nodes, their direction and their trust states as shown in Table 2. Node identity can be represent in two bytes

Table 2. Neighbor list table at sensor node.

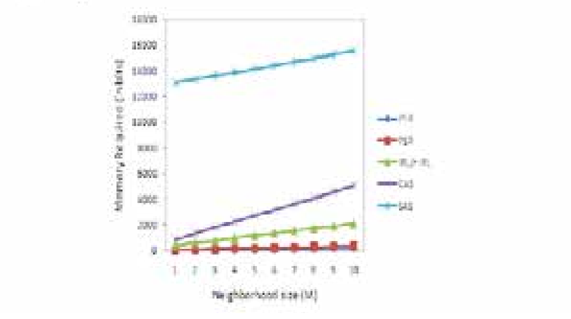
Neighbor nodeID (Integer)	Direction	Past interactions based on time window						Trust value
		Successful interactions ($S_{x,y}$)			Unsuccessful interactions ($U_{x,y}$)			
1	F (00)	10	...	5	4	...	1	90
2	B_R (01)	2	...	4	8	...	2	25
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
M	B_L(11)	5	...	7	0	..	3	70

Table 3. Memory requirement in bits.

PFR [3]	$(16+1)M$ bits
PSR [4]	$(16+16+1)M$ bits
SAS [5]	$K(4M+2N)+16M$ bits
CAS [5]	$K(6+7M)+16M$ bits
IRL / r-IRL	$M(26+32\Delta t)+k_{bs}^{+}+k_{r,bs}$bits

4.3. Energy Consumption Analysis

Figure 5. Memory consumption analysis: N= 100; K=8 bytes; Δt = 5; k_{bs}^{+} = 20 bytes; k_{r-bs} = 8 bytes.



In this section, we will show the efficiency of our routing strategies with existing schemes. Energy is computed based on the communication overhead (including transmission and reception cost, path length) introduced by our proposed routing protocols and compared it with other existing schemes.

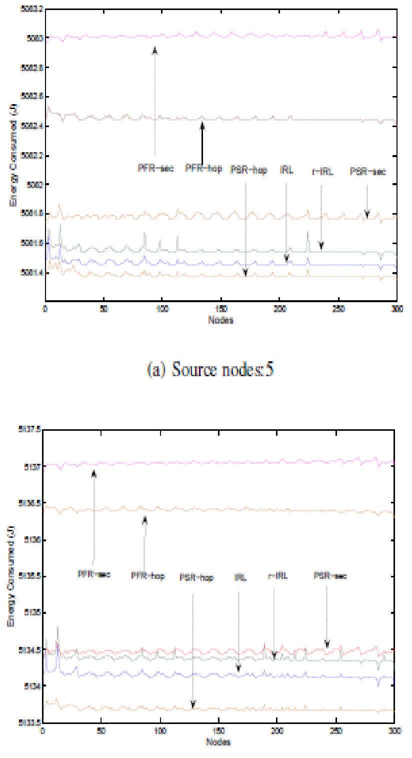
Table 4. Simulation parameters.

Network specific	Number of nodes	300
	Distance b/w nodes	50 units
	Mobility of nodes	zero
Node specific	Sensor node's Initial battery	1×10^6J
	Power consumption for trans.	1.6W
	Power consumption for recv.	1.2 W
	Idle power consumption	1.15W
	Carrier sense threshold	$3.65e^{-10}$W
	Receive power threshold	$1.55e^{-11}$W
	Frequency	$9.14e^8$
	Trans. & Recv. antenna gain	1.0
Protocol & Application specific	Application	CBR
	Reliability param. r for r-IRL	3
	h_{walk} param. for PFR & PSR	10

We have compared our proposed IRL and r-IRL algorithms with the four variations of phantom routing schemes that are:

1. Phantom single path routing scheme with hop-based approach (PSR-hop).
2. Phantom single path routing scheme with sector-based approach (PSR-sec).
3. Phantom flood routing scheme with hop-based approach (PFR-hop).
4. Phantom flood routing scheme with sector-based approach (PFR-sec).We did not compared our schemes with the SAS and CAS [5] schemes because the authors did not propose any routing strategy.

Figure 6. Energy consumption analysis: simulation time: 5,000.



(a) Source nodes:5

(b) Source nodes:10

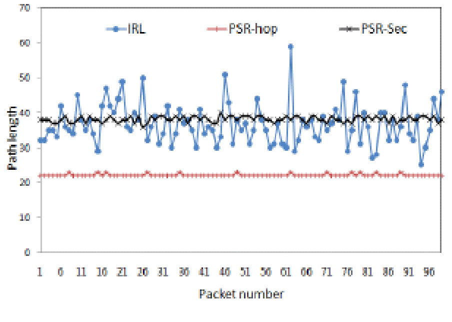
5.4. Path Diversity Analysis

Strength of route privacy is dependent on path diversity. High path diversity provides strong route privacy and low path diversity provides weak rout privacy. Path diversity can be categorized into two types.

1. Length variation: Path could be long or short and mainly dependent on routing scheme.
2. Path variation: Each packet may follow different route. It is also dependent on routing strategy..

Our proposed routing strategies (IRL and r-IRL) have both features. Because of the concept of*direction* , proposed schemes provide more length variation and because of the *randomness*

Figure 7. Path diversity of privacy schemes.



4.5. Discussion

From the memory, energy and path diversity analysis, we see that our solution is optimal especially with respect to the PSR-hop scheme. However, at a modest cost of memory and energy, our solutions provide full network level privacy as compared with the other existing schemes. This cost is justifiable because we have additionally achieved trustworthiness and reliability (in terms of packet reach-ability). With this level of resource consumption, our solutions can easily be used on real sensor nodes, for example, MICA2 sensor node has ATMega 128L micro controller (8 MHz @ 8 MIPS), 128 Kbyte program flash memory, 512 Kbyte measurement (serial) flash, and 4 Kbyte EEPROM .

5. Conclusions and Future work

Existing privacy schemes of WSNs only provides partial network level privacy. Providing full network level privacy is a critical and challenging issue due to the constraints imposed by the sensor nodes (e.g., energy, memory and computation power), sensor network (e.g., mobility and topology) and QoS issues (e.g., packet reach-ability and timeliness). Therefore, in this paper we proposed the first full network level privacy solution that is composed of two new identity, route and location privacy algorithms and data privacy mechanism. Our solutions provide additional trustworthiness and reliability at modest cost of energy and memory. We also proved analytically that our solutions provides protection against an adversary who is capable of performing privacy disclosure attacks such as eavesdropping and hop-by-hop trace backing.

In our future work, we will evaluate our proposed schemes from the perspective of computation cost that is required to perform encryption and random number generation.

REFERENCE

1. Xi, Y.; Schwiebert, L.; Shi, W. Preserving Source Location Privacy in Monitoring-Based Wireless | Sensor Networks. In Proceedings of Parallel and Distributed Processing Symposium (IPDPS | 2006), Rhodes Island, Greece, 2006. | 2. Habitat monitoring on Great Duck Island (Maine, USA), 2002. Available online: http://ucberkeley. citris-uc.org/research/projects/great duck island (accessed on 21 August, 2009). | 3. Ozturk, C.; Zhang, Y.; Trappe,W. Source-Location Privacy in Energy-Constrained Sensor Network | Routing. In Proceedings of the 2nd ACM workshop on Security of Ad hoc and Sensor Networks Washington, DC, WA, USA, 2004; pp. 88–93. | 4. Kamat, P.; Zhang, Y.; Trappe, W.; Ozturk, C. Enhancing Source-Location Privacy in Sensor Network Routing. In Proceedings of the 25th IEEE International conference on Distributed | Computing Systems, Columbus, OH, USA, 2005; pp. 599–608. | 5. Misra, S.; Xue, G. Efficient Anonymity Schemes for Clustered Wireless Sensor Networks. Int. J.Sens. Netw. 2006, 1, 50–63.