Location Privacy in Wireless Sensor Networks

Javier Lopez

Network, Information, and Computer Security (NICS) Lab
University of Malaga
Spain

https://www.nics.uma.es/jlm

NICS

FOSAD 2013

NESSoS

Agenda

- WSN basics
- Privacy in WSN
  - Suitability of Existing Approaches
- Privacy of Location
  - Node Anonymity
  - Source-Location Privacy
    - Local, Global and Internal adversaries
  - Receiver-Location Privacy
    - Local, Global adversaries
  - Anonymous Topology Discovery
- Final Remarks

NICS

2

WSN BASICS

3

Introduction to WSNs

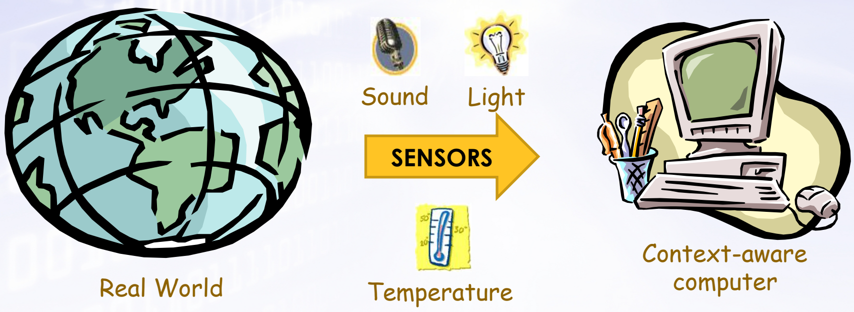
- Humans are able to feel the world thanks to our senses

NICS

4

Introduction to WSNs

- Sensors are to computers what senses are to humans



NICS

From sensors to sensor nodes

Autonomous Computer

+

Sensing board

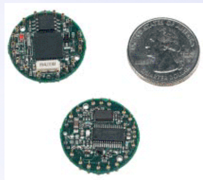
=

SENSOR NODE

NICS

Commercial Products

- The market already offers a number of sensor network hardware products
  - not only for research purposes,
  - but for the integration and deployment in real-world ubiquitous applications:
    - EMS nodes by *Sensicast Systems*,
    - EM chips by *Ember Corporation*,
    - Mesh485 by *Millenial Net*,
    - Mote kits by *Crossbow Technology*,
    - SmartMesh-XR by *Dust Networks*,
    - Tmote Invent System by *Moteiv*.
    - etc.



MICADOT motes

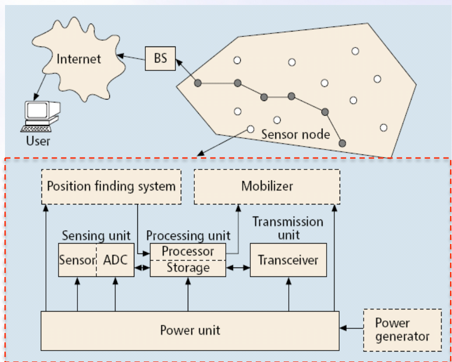
Telos mote

NICS

7

From sensor nodes to WSN

- WSNs are ad-hoc networks comprised of [Akyi02] :
  - Sensor Nodes are battery-powered devices with limited capabilities
    - Monitor the environment, and transmit these data to nearby nodes
    - Operate and cooperate in adhoc manner using radio interfaces
    - Support multiple communication paths
    - Provide routing capabilities
  - Base station (sink)
    - Has no limited resources
    - Collect and process data received from sensor nodes



NICS

Limitations

- For the case of *Mica* family (*Mica2*, *Mica2dot*, *MicaZ)*, and *Telos* nodes:
  - Processor:
    - 8-bit Atmel ATmega processor
    - Telos: 16-bit TI MSP430 processor
  - Memory:
    - 128 KB ROM and 4 KB RAM
    - Telos: 48 KB ROM and 10 KB RAM
  - Speed:
    - Mica2dot: 4 MHz
    - Mica2 and MicaZ: 7.37 MHz
    - Telos: 8MHz

NICS

Limitations

- Communications:
  - Mica2dot and Mica2 deliver up to 20 kbps on a single shared channel, with a range of up to around a few hundred meters
  - MicaZ and Telos deliver up to 250 kbps.
- Software:
  - *TinyOS* operating system
    - Highly optimized (small, fast,…)
    - Support real-time tasks (multi-threaded, events-oriented)
  - C variant called *nesC* for programming purposes
    - featuring an event-driven concurrency model

NICS

Limitations

	Btnode 3	mica2	mica2dot	micaz	telos A	tmote sky	EYES
Manufacturer	Art of Technology	Crossbow			Imote iv		Univ. of Twente
Microcontroller	Atmel Atmega 128L				Texas Instruments MSP430		
Clock frequency	7.37 Mhz		4 MHz	7.37 MHz	8 MHz		5 MHz
RAM (KB)	64 + 180	4	4	4	2	10	2
ROM (KB)	128	128	128	128	60	48	60
Storage (KB)	4	512	512	512	256	1024	4
Radio	Chipcon CC1000 315/433/868/916 MHz 38.4 Kbauds			Chipcon CC2420 2.4 GHz 250Kbps IEEE 802.15.4			RFM TR1001868 MHz 57.6 Kbps
Max Range (m)	150-300			75-100			
Power	2 AA batteries		Coin cell	2 AA Batteries			
PC connector	Through PC-connected programming board				USB		Serial Port
OS	Nut/OS	TinyOS					PEEROS
Transducers	On acquisition board				On board		On acquisition board
Extras	+ Bluetooth radio						

NICS

WSN Applications

- Generally speaking, WSNs can be used in applications where sensors are unobtrusively embedded into systems, consequently involving operations like:
  - Monitoring
  - Tracking
  - Detecting
  - Collecting
  - Reporting

- By sectors, WSNs can be used in:
  - Agricultural
  - Business
  - Critical infrastructure protection
  - Environment
  - Health care
  - Homeland security
  - Industrial
  - Military applications
  - etc.

NICS

WSN Applications

- Specific applications:
  - farmland monitoring
  - animal identification and tracking
  - cultivation conditions (temperature, humidity, etc.)
  - inventory control
  - goods tracking and delivery
  - smart office
  - supply of water and electricity
  - freeway traffic monitoring and control
  - detection of structural integrity
  - problems in buildings
  - wildlife habitat monitoring
  - microclimate control
  - detection of out-of-tolerance
  - environmental conditions
  - recording wild animal habits
  - emergency medical care
  - remote medical monitoring
  - medicines tracking
  - frontiers surveillance
  - detection of illegal materials in
  - custom controls
  - monitoring factory instrumentation
  - remote control of manufacturing
  - systems
  - collecting pollution levels
  - detection of structures vibrations
  - target tracking
  - detection of biological or chemical
  - weapons
  - location of vehicles and arms
  - wearable smart uniforms
  - etc.

NICS

13

WSN Applications...for Internet

- Still a wide range of applications to come when sensors can – directly – exchange information with entities on the Internet:
  - reaching, for instance, home environments.
  - creating what already has been called:
    - "network of things"
    - "Internet of things"
    - "tangible Internet"
    - "Internet of objects"
    - "Internet of Everything"
    - etc.

NICS

14

WSN Communication Architecture

15

WSN Communication Architecture

- Sensors operate and cooperate in an ad hoc manner using radio interfaces, resulting in a mesh architecture where nodes:
  - communicate directly only with nodes nearby due to limited power
    - some nodes communicate with a base station
  - support multiple communication paths
  - provide routing capabilities

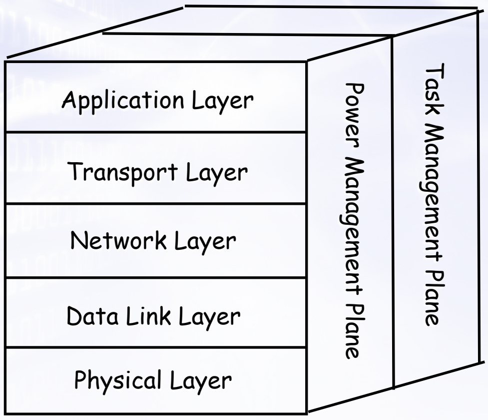
what turns out to be an advantage in comparison with 802.11 and Bluetooth.

NICS

16

WSN Communication Architecture

- The communication architecture may be initially considered in the following hierarchical way

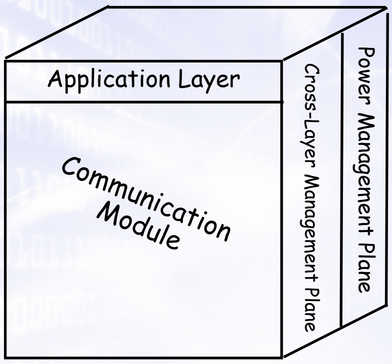


NICS

17

WSN Communication Architecture

- Due to cross-layer melting, it is evolving to the following:



  - Cross-layer contributes to autonomy and self-configuration of the nodes because any component can directly access to resources and processes provided by another component
  - Flexible access to information and control is convenient due to: (i) inherent limitations of sensors, (ii) specific applications requirements

NICS

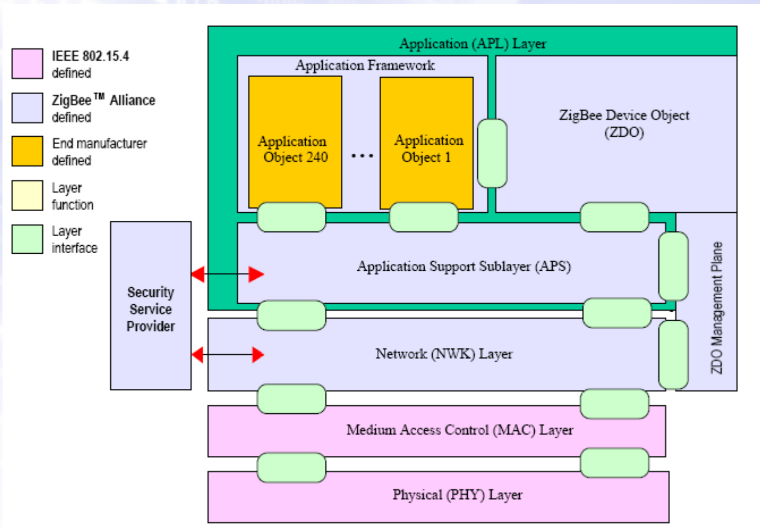
18

WSN Communication Architecture
The case of Zigbee

- ZigBee: Specification for WSN
  - Built upon IEEE 802.15.4
    - Standard for WPAN
    - Low energy consumption, low transmission rate (250kbps), low cost
  - Security: AES-128
- Hierarchical model
- But with limited support to cross-layer
  - Management
  - Security

NICS

WSN Communication Architecture
The case of Zigbee



NICS

PRIVACY IN WSN

21

Threats to WSNs

- Due to the resource limitation of sensor nodes, WSNs are highly vulnerable to threats and attacks [Walt07]
  - Information flow attacks
    - Eavesdropping, modification, reply attacks
  - Denial of Service
    - Jamming, network flooding, battery exhaustion
  - Physical attacks
    - Node destruction, node compromise
  - Node impersonation
    - Node Replication, Sybil attack
  - Specific attacks
    - Wormhole, sinkhole, selective forwarding
- However, among the threats to WSN, privacy concerns on information being collected and transmitted have received less attention

NICS

22

Originality of privacy in WSN

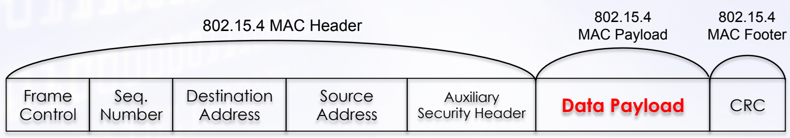
- Privacy protection has been extensively studied in wired and wireless networking, and a number of techniques have been designed
  - But, as shown later, they can not be directly applied to WSN because these networks have special features:
    - Sensor-node resource constraints (some cryptographic techniques can not be executed)
    - Conflict of anonymity requirements/goals (traditional traffic analysis countermeasures are not always useful)
    - Uncontrollable environment (physical attacks plus key retrieval)
    - Topological constraints (multiple hops scheme with unbalanced traffic loads)

NICS

23

Privacy of payload information

- In WSNs, the private information to protect, in principle, would be that one included in the packets transmitted
  - Payload information: data collected by a sensor and transmitted to a server
- That information traversing the network can be protected from eavesdropping
  - by using some of the traditional confidentiality and integrity mechanisms.



NICS

24

Privacy of contextual information

- However, even if the payload data is encrypted, the attacker can still attack in another way
- That is, by observing and analyzing the communications, an attacker could retrieve contextual information (what is also private data)
  - about the network itself
  - and about the type of data being collected by the WSN
  - not only the occurrence of an event must be protected; also the moment in time when the event takes place: temporal privacy
    - if an adversary is able to make an association between the time and position of the events being monitored, he will be able to predict future behaviours.

NICS

25

Privacy of contextual information

- What information can be learnt by the attacker in this way? Simple observation of network traffic can reveal a lot [Pai08]
  - Frequency range can be used to determine
    - Type of sensor
      - Exploit specific platform vulnerabilities
    - Owner of the network
      - Different organizations are designated different frequency bands
  - Transmission rate can provide information about
    - Amount and nature of events
      - The presence of events triggers message transmission
    - Distance to the sender
      - Time of arrival of packets can be used to calculate the distance to the sender

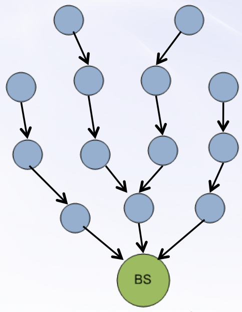
Commonly used name	Frequency range (MHz)
Mica or Mica1	902 to 928 433.1 to 434.9
Mica2	868 to 870 902 to 928 433.1 to 434.8 313.9 to 316.1
Mica2Dot	868 to 870 902 to 928 433.1 to 434.8 313.9 to 316.1
Micaz	2400 to 2483.5
Cricket	433.1 to 434.8
IRIS	2400 to 2483.5
TelosB	2400 to 2483.5

NICS

26

Privacy of contextual information

- Packet size might reveal the
  - Proximity to the base station
    - Due to certain data-aggregation mechanisms the closer to the base station the larger the packet might be
  - Type and precision of the data collected
    - Complex data types need larger payloads
- Routing protocols give information about
  - Network topology
    - Messages are sent to the base station, and packets usually follow a pre-fixed route to its destination



NICS

27

Privacy of contextual information

- In summary, even when the payload is encrypted, there is a lot of things that the attacker can learn by observing and analyzing the flows of information (traffic) in the WSN
- Could we use privacy solutions already developed for the Internet and its applications?

NICS

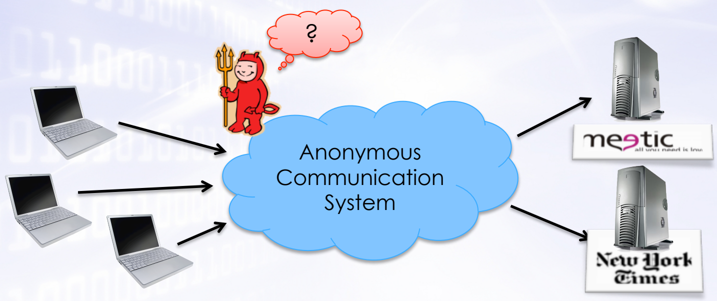
28

Suitability of Existing Approaches

29

Internet Anonymous Communication Systems

- Anonymous communication systems (ACS) were devised to prevent traffic analysis attacks in Internet applications
- The question is: are these ACS suitable for WSNs?



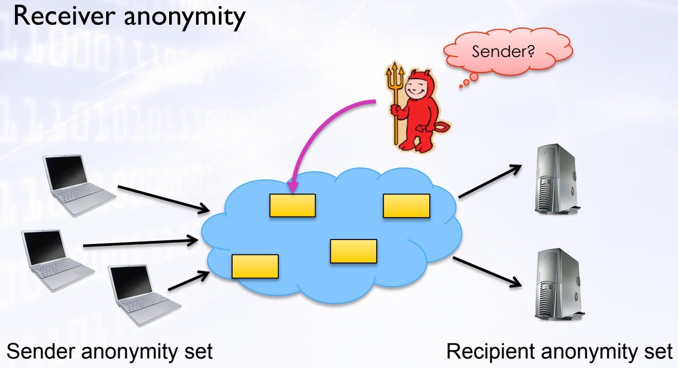
- ACSs focus on different aspects depending on the requirements of the user [Pfit10]

NICS

30

Anonymity Properties

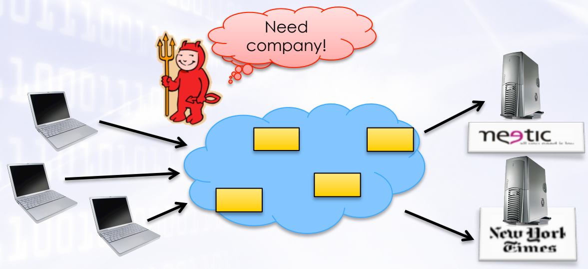
- Anonymity
  - An attacker cannot sufficiently identify a subject within a set of subjects (anonymity set) with potentially the same attributes
    - Sender anonymity
    - Receiver anonymity



NICS

Anonymity Properties

- Unlinkability
  - An attacker cannot sufficiently distinguish whether two or more items of interest (IoI) are related or not

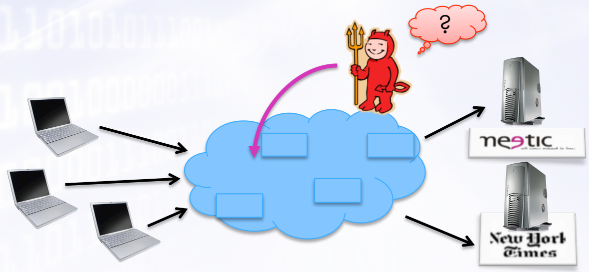


  - Relationship unlinkability hides the correspondence between a user and the servers being accessed

NICS

Anonymity Properties

- Undetectability
  - An attacker cannot sufficiently distinguish whether an item of interest exists or not
  - It aims to protect the items of interest as such
    - A steganographic message passes unnoticed to attackers
    - Dummy traffic also hides the presence of real traffic



NICS

33

Anonymity Properties

- Unobservability
  - This concept implies both undetectability of the IoI and anonymity of the subjects involved in that IoI
  - Even if a subject could detect an IoI, the other subjects involved in the IoI remain anonymous
    - Sender unobservability
    - Recipient unobservability
    - Sender-Recipient unobservability

NICS

34

Anonymity Mechanisms

- Anonymity properties have been developed in ACS by combining different techniques:
  - Symmetric/Public-key crypto
  - Layered encryption
  - Packet delay/replay/injection
  - Multicast/Broadcast communications

	Main goal	**Architecture**	**Techniques**						
			SK	PK	LE	PD	PR	FT	MB
Single-proxy	Sender Anonymity	Centralized	√						
Mix-nets	Unlinkability			√	√	√			
Onion routing			√	√	√			√	
Tor			√	√	√				
Crowds	Sender Anonymity	Decentralized	√						
Hordes			√	√					√
GAP	Unobservability		√	√		√	√	√	
DC-nets			√						√
Herbivore			√	√					√

NICS

35

Example 1: Mix-nets

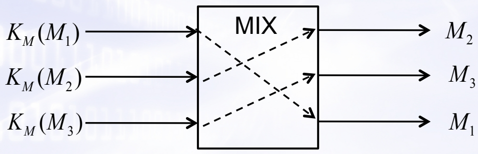
- Mix-nets [Chau81] are composed of a set of devices which are place in between senders and recipients
- Mixes are based on
  - Message delay
  - Public-key crypto
- Attacker model
  - Eavesdroppers
  - Can also provide sender anonymity w.r.t. recipient
- Not intended for real-time applications
  - Originally designed for mailing systems

NICS

36

Example 1: Mix-nets

- Store-and-forward device that randomly permutes and decrypts inputs
  - Messages are output as re-ordered batches



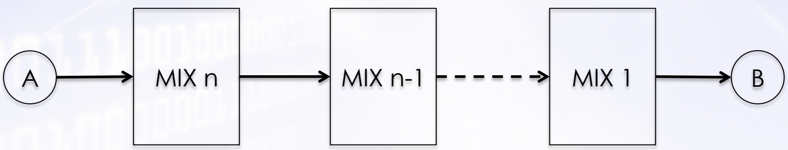
- An adversary can't correlate inputs and outputs because of temporal storage and decryption of messages

NICS

37

Example 1: Mix-nets

- Communicating through a single mix might not be sufficiently secure
  - A single mix knows both sender and destination
- The user selects a series of mixes and creates a layer of encryption for every mix
  - Every mix only knows its predecessor and successor in the path



- A single honest mix prevents input-output correlation

NICS

38

Example 1: Mix-nets

- The implementation of mixnets over WSNs present several limitations
  - Every source node is required to
    - Perform N + 1 public-key operations per transmitted packet
    - Have global network knowledge to be able to determine the transmission path
  - Every intermediate node is required
    - Perform 1 public-key operation per received packet
    - Temporarily store a large number of packets
    - Message padding is required for message indistinguishability
    - Output a single re-ordered batch of messages
    - Nodes in the vicinity of the base station have even higher traffic rates
  - Many WSN applications require real-time monitoring

NICS 39

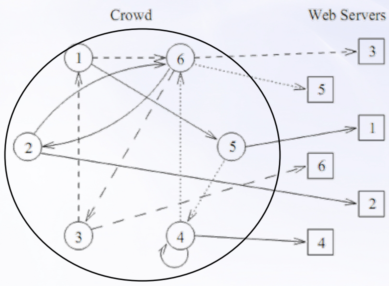
Example 2: Crowds

- Crowds [Reit98] is a decentralized solution where a set of users collaborate to perform requests to servers on behalf of its members
- Crowds are based on
  - Symmetric-key crypto
  - Random intermediate node selection
  - No Public-key crypto, dummy traffic nor padding!
- Attacker model
  - Local (and static) observers
  - Colluding (internal) members
  - No protection against global observers!
- Intended for near real-time applications
  - Originally designed for web browsing

NICS 40

Example 2: Crowds

- The Crowd consists of a dynamic collection of users controlled by *the blender*
  - The blender is in charge of the crowd admittance process
- Members initiate requests to various servers by creating a random path within the crowd
  - The request is finally submitted by a random member
  - Subsequent requests and replies follow the same path
  - Packets belonging to a path are identified by a changing path_id



NICS

41

Example 2: Crowds

- Local eavesdroppers are static and observe inputs/outputs from a single node
  - May recognize the initiator and destination only if observes the right member
  - Probability decreases with the crowd size
- End servers cannot determine the initiator
  - The initiator never submits the packet to the server in the first step
  - All members are equally probable to be the initiator
- Colluding members might want to know the initiator
  - Suspect from the member that immediately precedes the first collaborator in the path
  - Static paths reduce the probability of this type of attacks

NICS

42

Example 2: Crowds

- The potential application of the Crowds model to WSNs is restricted by:
  - High memory requirements
    - Path_id translation table
    - N – 1 shared keys (1 key per member)
  - Limited number and complexity of the operations
    - 1 Symmetric-key operation per packet
    - 1 Path_id replacement per packet
  - Weak adversarial model
    - Static attackers have a very limited success probability
  - Different requirements
    - Source anonymity with respect to the sink is counterproductive in WSNs

NICS

43

Originality of privacy in WSN

- The high overhead of traditional solutions is not the only limiting factor to the application of ACSs to WSNs [Rios2012]

- Most traditional anonymity solutions aim to hide the relationship between senders and receivers (unlinkability)
  - Unlinkability is not necessary in WSNs because the model of communication is known (nodes-to-sink)
  - The attacker already knows that any sensor node will communicate with the base station

NICS

44

Originality of privacy in WSN

- Some ACSs provide the users the opportunity to hide their identity to the server (anonymity)
  - Providing source anonymity with respect to the sink is detrimental for the normal operation of the network
  - A proper manage and control of the environment being monitored requires the sink to be aware of the data sender
  - However, in WSNs source anonymity is indeed important against external observers or compromised intermediaries

NICS

45

Originality of privacy in WSN

- Also, while in the Internet it is not totally necessary to hide the participation of senders or receivers, this is an essential requirement in WSN

- In WSNs it is required to hide the presence of event messages (i.e. event undetectability)

NICS

46

Originality of privacy in WSN

- In summary, high overhead introduced by Internet solutions is a limiting factor to their use in WSNs
  - Sensor-node resource constraints (some cryptographic techniques can not be executed, battery powered, limited memory)
- Additionally, the usual properties provided by those solutions are not always suitable in WSNs
- Hence new tailored solutions must be designed for WSNs

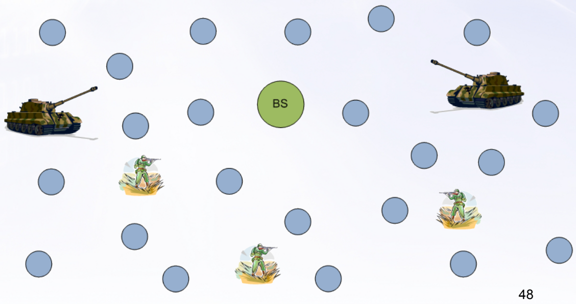
Property	Traditional Solution	WSN
Unlinkability	Observers try to know with whom a user communicates	All sensors are known to send data to the sink ✗
Sender Anonymity	Servers might try to profile or track their users	The data source needs to be known by the sink ✗
Unobservability/ Undetectability	Users might be reluctant to show their participation in the system	Hiding the presence of senders hides the presence of events

NICS

47

Originality of privacy in WSN

- One additional reason:
  - Attacker's goal is another reason that makes WSN scenarios a special privacy case
    - It is of paramount importance to find the source of the events
    - Also very important to find the base station
- Because of this, most of research so far has focused on Location Privacy



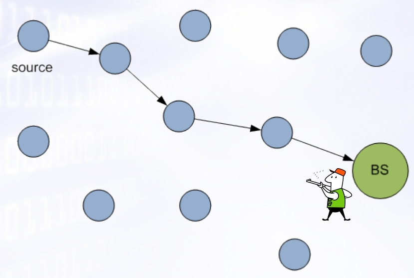
NICS

48

PRIVACY OF LOCATION

Objective

- The objective of location privacy is to prevent an attacker from determining the location of specific nodes of interest to him



NICS

Traffic analysis attacks

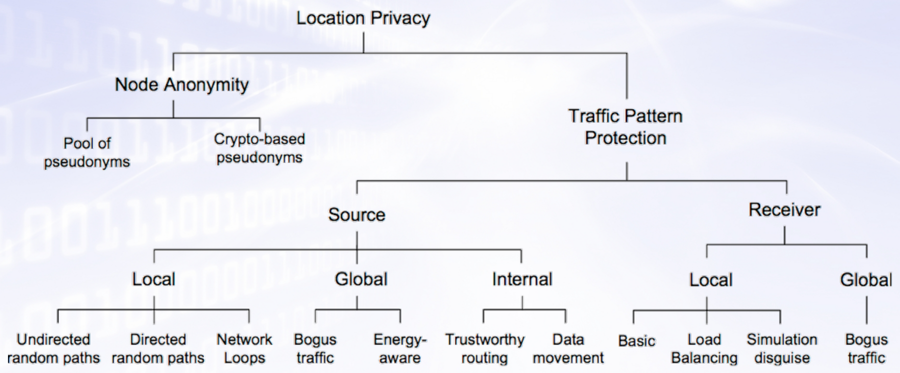
- Different adversarial models can be found according to the attacker's ability to:
  - Disturb network operation
    - Passive: simply eavesdrops and performs traffic analysis attacks
    - Active: can also create, modify or inject packets, destroy nodes, ...
  - Compromise nodes
    - External: has no knowledge about the internals of the node
    - Internal: is able to compromise nodes, access cryptographic material and algorithms
  - Observe communications
    - Local: has monitoring radius similar to a sensor node
    - Global: has the ability to capture all the traffic generated by the network

NICS

51

Classification of protection mechanisms

- We classify the protection mechanisms depending on the asset to be protected and the attacker's capabilities [Rios11a]



NICS

52

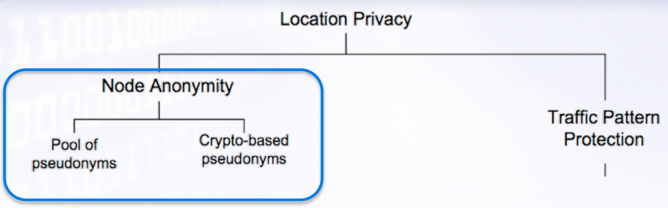
Traffic analysis attacks

- Hence, the power of the adversary will determine the types of attacks he might perform. Typical attacks are [Shao08]
  - Content analysis attack
    - Examine the content of an event message to determine if the location of the node is contained in plaintext in the payload or headers
  - Traceback attack
    - An attacker equipped with a directional antenna can estimate the angle of arrival of the signal and arrive at the immediate sender of a message
  - Rate monitoring attack
    - The number of messages being sent by the nodes can be used to determine the location of (or direction to) the important nodes
  - Time correlation attack
    - The observation of the transmission times between a node and its neighbours the attacker may deduce the transmission path

NICS

53

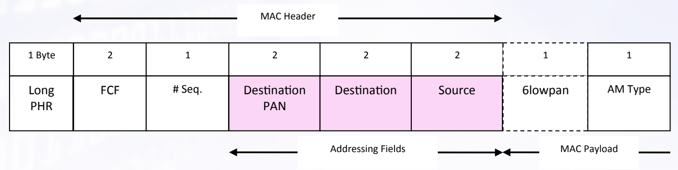
Node Anonymity



54

Node Anonymity

- As previously mentioned, the first step to protect location privacy is to encrypt packet contents
  - that is, hide any information that might be used by the attacker to learn the nodes involved in the communication
- But there is also information contained in the packet headers that is usually not protected: identifiers of sender and recipient.
  - Often, the identifier of a node is enough to determine its location



TinyOS 2.x MAC Header

NICS

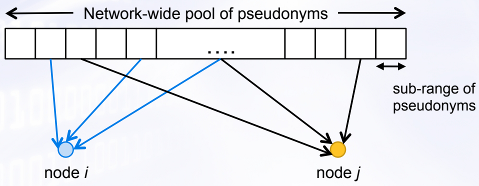
Pseudonyms

- A pseudonym is a name or identifier that can be used instead of a real name
  - Pseudonyms are used to protect the real identities of the nodes
- Using fixed pseudonyms eventually provides no protection because the attacker relates a pseudonym with a node
- Several schemes have been proposed to create dynamic pseudonyms
  - Simple Anonymity Scheme (SAS)
  - Cryptographic Anonymity Scheme (CAS)
  - Hashing-based ID Randomization (HIR)
  - Reverse HIR (RHIR)

NICS

Pool of Pseudonyms: SAS

- Simple Anonymity Scheme (SAS) [Misr06]
  - Pre-deployment phase:
    - Defines a K-bit pseudonym address space
    - Each sensor is assigned with N randomly distributed sub-ranges of l bits
    - The BS stores the pseudonym ranges for each node in order to figure out the correct decryption key



    - Some sub-ranges can be left "free" for future use in case it is necessary to revoke some neighbours

57

Pool of Pseudonyms: SAS

- Simple Anonymity Scheme (SAS)
  - Post-deployment phase:
    - Every node randomly assigns one sub-range to each of its neighbours
    - The sub-ranges to be used are securely exchanged
    - Each node builds a pseudonyms table to map pseudonym to and from its neighbours together with the corresponding shared key

Index TX	Sub-Range TX	Sub-range RX	Index RX	Shared key
…	…	…	…	…
Ind_y	$IDxy^{ini}$, $IDxy^{end}$	$IDyx^{ini}$, $IDyx^{end}$	Ind_x	Kxy
…	…	…	…	..

    - Node X generates a sender ID and receiver ID for every message, as
      - SenderID = Ind_y \|\| random($IDxy^{ini}$, $IDxy^{end}$)
      - ReceiverID = Ind_x \|\| random($IDyx^{ini}$, $IDyx^{end}$)
    - Node Y uses Ind_y to search for pseudonym in its table

58

Pool of Pseudonyms: SAS

- SAS presents some limitations with respect to the memory requirements
  - Every node needs to store a number of sub-ranges of pseudonyms
  - The pool of pseudonyms will eventually be exhausted
  - Some sub-ranges may be exhausted while others may be not
- No limitations from a computational point of view
  - It is only necessary to find the Index in the pseudonyms table
  - Check the received pseudonym is in range, and
  - Decrypt with the shared key
- When a node is compromised, the attacker obtains all the pseudonyms and shared secrets
  - Revocation of nodes is useful but limits the pool of pseudonyms
- An outsider could determine the ranges of pseudonyms used by a particular node or even impersonate it

NICS

59

Crypto-based Pseudonyms: CAS

- Cryptographic Anonymity Scheme (CAS) [Misr06]
  - Uses Keyed Hash Functions (KHF) to generate pseudonyms
  - Pre-deployment
    - Nodes are assigned a pseudo-random function f_x, a key shared with the BS K_{BSx}, and a random seed s_{BSx} for communication with the BS
    - Every pair of neighbours share a key K_{xy} and random seed s_{xy}
    - Every node builds a pseudonym table with an entry for each neighbour

Index	Seed	# sequence	Shared key	Neigh Index
…	…	…	…	…
Ind_x	s_{xy}	seq_{xy}	K_{xy}	Ind_y
…	…	…	…	…

    - The sequence number is used during message generation for the creation of indistinguishable pseudonyms

NICS

60

Crypto-based Pseudonyms: CAS

- Cryptographic Anonymity Scheme (CAS) [Misr06]
  - Communication phase
    - Messages from x to *BS* (going through node y) have the following form

SID || RID || EncryptedPayload || seq_{xy}

- SID = Indx || $H_{KBSx}(s_{BSx}$ XOR $seq_{xy})$
- RID = Indy || $H_{Kxy}(s_{xy}$ XOR $seq_{xy})$

    - Again, the recipient uses the Index to find s_{xy} and K_{xy} in its table and check the validity of RID
    - The SID is used by the BS to check the source of the message and obtain the decryption key

NICS

61

Crypto-based Pseudonyms: CAS

- CAS is more computationally intensive than SAS but it reduces the memory requirements
  - The source needs to generate 2 pseudonyms
  - Any intermediary generates 1 pseudonym
  - A non-intended recipient also need to compute the hash value to check whether the RID is intended to it
  - External attackers learn nothing by observing the pseudonyms

- Both SAS and CAS are based on the assumption that an attacker cannot compromise the secrets shared between nodes

NICS

62

Crypto-based Pseudonyms: Hash Chains

- To reduce the impact of shared secrets being compromised, Keyed Hash Chains are used to generate pseudonyms [Ouya07]

$$ID \rightarrow H_K(ID) \rightarrow H_K(H_K(ID)) \rightarrow \ldots \rightarrow H_K^n(ID)$$

- A sensor node can delete its previous ID and generate a new one after sending a message
  – Provides backward anonymity as the hash function cannot be reversed

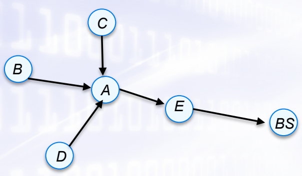
$$ID \nleftarrow H_K(ID) \nleftarrow H_K(H_K(ID)) \nleftarrow \ldots \nleftarrow H_K^n(ID)$$

- They proposed two schemes
  – HIR (Hashing-based ID Randomization)
  – RHIR (Reverse HIR)

63

Crypto-based Pseudonyms: HIR

- Hashing-based ID Randomization (HIR)
  – Sensors determine their uplink and downlink neighbours and share pairwise keys with them
  – Create a table that includes the keyed hash values of their neighbours



Depth	Neigh Hashed ID	Link
1	$H_{KAB}(ID_A)$	down
2	$H^2_{KAC}(ID_A)$	down
4	$H^4_{KAD}(ID_A)$	down
2	$H^2_{KAE}(ID_A)$	up

  – Messages have the following form: $M = H_R \| H_S \| t \| Data$
    - $H_R = H^{depth_Y}_{K_{XY}}(ID_Y)$ pseudonym used to identify the next recipient of M
    - $H_S = H^t_{KX}(ID_X)$ pseudonym used for the BS to identify the original source (t indicates the depth of H_S)

64

Crypto-based Pseudonyms: RHIR

- Reverse Hashing ID Randomization (RHIR)
  - RHIR uses the hash chain in reverse order
  - The sensor node needs to compute the hash chain first and store it locally

generation →

$$ID, H_K(ID), H_K^2(ID), \ldots, H_K^n(ID)$$

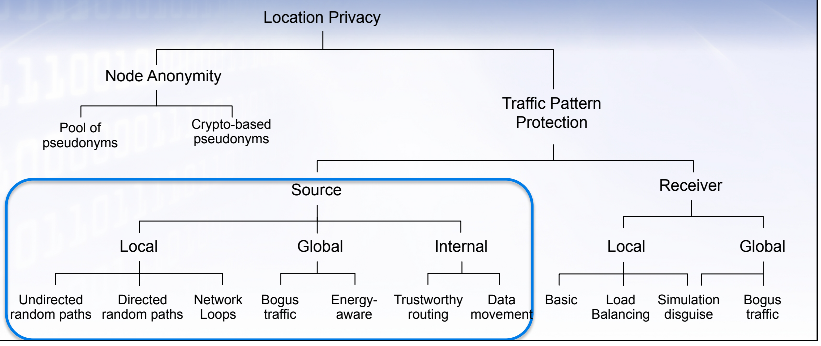
← use

  - An attacker cannot obtain the next pseudonyms to be used even if he compromises the key *K*
  - However, it limits the number of available pseudonyms
    - Generating a large number of pseudonyms implies a large memory consumption

NICS

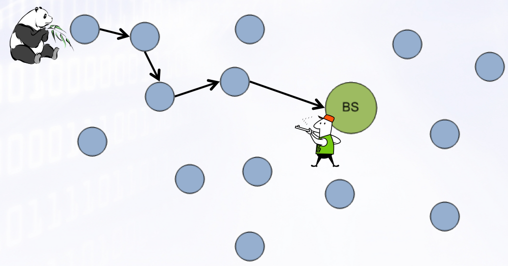
65

Source Location Privacy (SLP)



Source Location Privacy (SLP)

- Aims to protect the location of nodes generating event messages [Oztu04,Kama05]
  - The location of the source nodes indicates the location of events
- Problem motivated by the Panda Hunter Game:



NICS

67

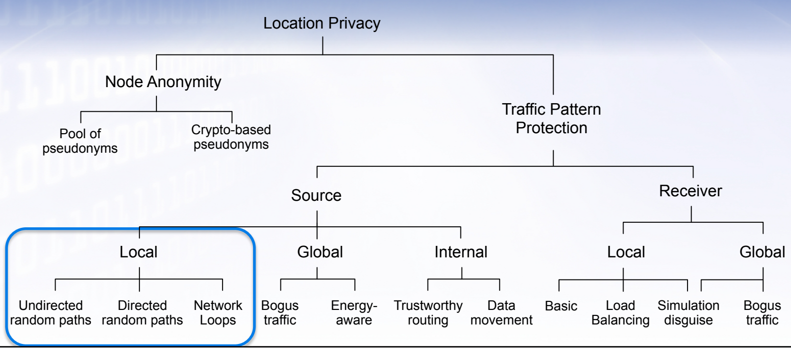
Source Location Privacy (SLP)

- Panda vs Hunter:
  - Sensor nodes report the presence of the panda as soon as they sense it
  - Messages are sent in a hop-by-hop manner towards the base station
  - The hunter is equipped with a device that allows him to listen to the communications generated by sensor nodes
  - Encrypting the content of the messages cannot help because the mere existence of messages is indicative of the occurrence of events
- How to provide a solution depends on the model of attacker:
  - local
  - global
  - …

NICS

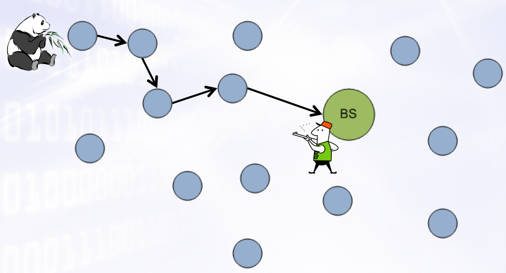
68

SLP: Local Adversaries



Local Adversaries

- The attacker usually stays close to the base station [Oztu04,Kama05]. Upon the reception of a packet he will jump in that direction.
  – The process is repeated for each received packet



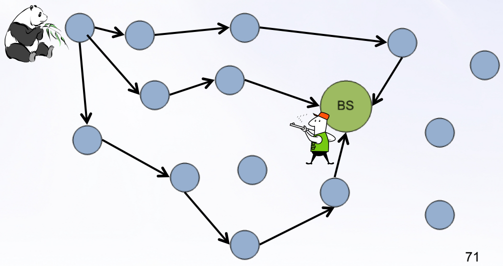
- The attacker finds the source because messages always follow the same route to the base station

NICS

70

Local Adversaries: Solutions

- The goal is to mislead the adversary in order to increase the safety period
  - which is the number of packets sent by the source before the panda is caught (time the panda is safe)
- Most of the proposed solutions to counter local adversaries are based on the randomization of routes
- That is, using different paths for different packets could be an effective defence



NICS

71

Local Adversaries: Solutions

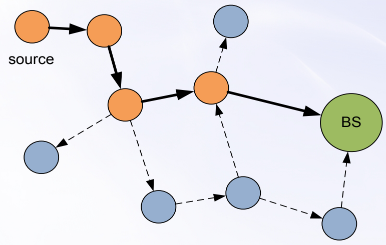
- However, the randomization of the routes has a cost:
  - it introduces some delay in the arrival of packets to the base station
  - possible increase in the probability of packet loss due to the use of longer paths
  - significant increase in energy consumption due to the increasing number of hops a packet needs to perform to reach destination
- Goal:
  - How to use different paths while avoiding aforementioned drawbacks?

NICS

72

Phantom Routing

- Phantom Routing [Oztu04, Kama05] was the first solution proposed, and most other solutions concentrate on improving it
- It was developed after analysing the privacy implications of widely used routing protocols in WSNs
  - Single-path / shortest-path routing
    - Shortest safety period
    - Lowest power consumption
  - Baseline flooding
    - Shortest safety period
    - Largest power consumption
  - Probabilistic flooding
    - Increased safety period
    - Reduced power consumption
    - Reduced delivery probability



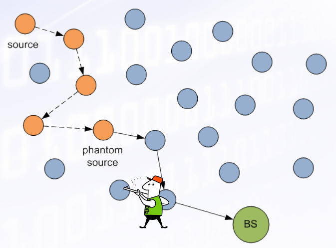
Baseline Flooding showing shortest path (in orange)

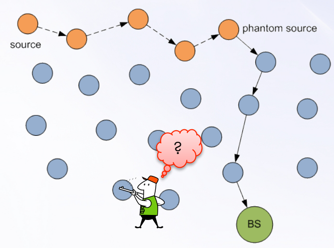
NICS

73

Phantom Routing

- It consist of two phases
  1. Random or directed walk
  2. Flooding or single path
- During the walking phase the packet travels for *h* hops until it reaches a random phantom source
  - The phantom source leads the adversary away from the real source
  - If no packets are received the attacker returns





NICS

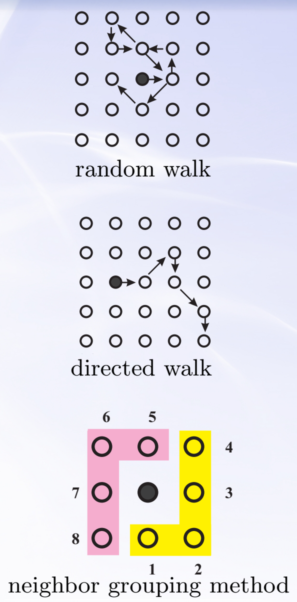
First transmission

Next transmission

74

Phantom Routing

- The walking phase must be carefully designed in order to avoid
  – Similar consecutive paths
  – Phantom sources close to the real source node
- The directed random walk aims to prevent previous problems by grouping neighbours into closer and further
- Main limitations of Phantom Routing
  – Increased latency and energy consumption



NICS

75

Local Adversaries: Solutions

- As mentioned, most other solutions concentrate on improving Phantom Routing. Among them, we highlight:
  – Undirected random walks (URW)
    - GROW (*Greedy Random Walk*)
    - Random Parallel routing
    - Cross-layer source location privacy
  – Directed random walks (DRW)
    - PRLA (*Phantom Routing based on location angle*)
    - WRS (*Weighted Random Stride*)
    - RRIN (*Random Intermediate Node*) & STaR (Sink Toroidal Region)
  – Network Loop methods (NL)
    - CEM (*Cyclic Entrapment Method*)
    - NMR (*Network Mixing Ring*)

NICS

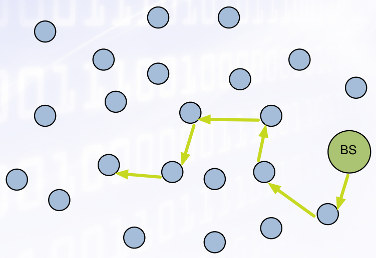
76

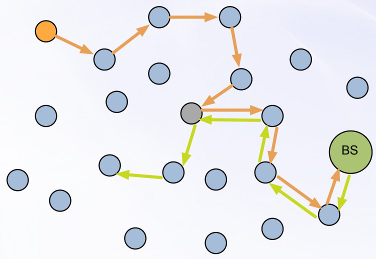
- Undirected random walks (URW)
- Directed random walks (DRW)
- Network Loop methods (NL)

NICS

URW: Greedy Random Walk

- GROW (*Greedy Random Walk*) is a greedy two-way random walk algorithm [Xi06]
  1. Base station sends a random walk message to build a path of receptors
  2. Every new message sent by the source will follow a random walk that will eventually hit the path of receptors

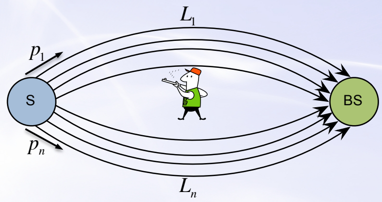




NICS

URW: Random Parallel Routing

- In [Wang09] every sensor node is pre-assigned *N* parallel paths to the base station
- Paths must be geographically separated so that attacker cannot overhear packets on other paths



- Messages must be evenly distributed on each path so that the attacker does not have an advantage by choosing a particular path

NICS

URW: Random Parallel Routing

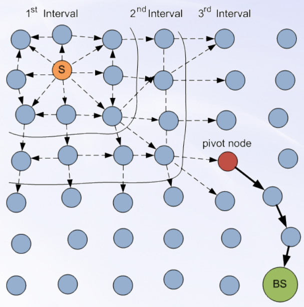
- There are some limitations in Random Parallel Routing with respect to:
  - Complexity
    - Path calculation is a complex task
    - Storing *N* paths requires much memory
    - The path to follow must be stored in each packet
  - Privacy
    - In practice, since paths are parallel, capturing few packets in a path help to infer the direction to the source

$$T <<< \sum_{i=1}^{n} L_i$$

NICS

URW: Cross-layer source location privacy

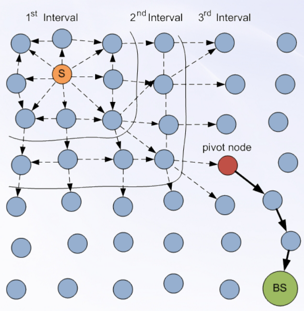
- Cross-layer source location privacy [Shao09a] benefits from beacon messages to conceal the walking phase of Phantom Routing
- Beacons are broadcasted periodically to announce node presence and for network configuration purposes
  - Beacons are transmitted regardless of the presence of events in the field
  - Contain a 15 bytes payload
  - Event data can be hidden (encrypted) within beacon messages without raising suspicion



81

URW: Cross-layer source location privacy

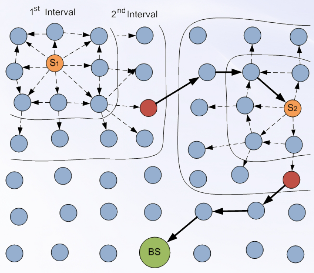
- Beacons travel for several hops to a pivot node (~phantom source), which passes the event data to the routing layer
- Provides perfect privacy for all attackers within the beaconing range as long as they are not within range of the pivot node or on the path from the pivot node to the BS
- Therefore, source nodes must choose different pivot nodes or the attacker will be able to reach the "edge" of the beaconing area
  - Specially when the distance between the source and pivot cannot be large because it has a significant impact on the delivery time



82

URW: Cross-layer source location privacy

- A double cross-layer solution improves privacy when the attacker is near the BS
  - The routing layer usually implements a shortest-path routing algorithm
- The pivot node does not send the packet to the BS directly
- The pivot node chooses a random node (using the routing layer) to start a second MAC-layer broadcast
- Additional phases increase latency and don't necessarily improve privacy

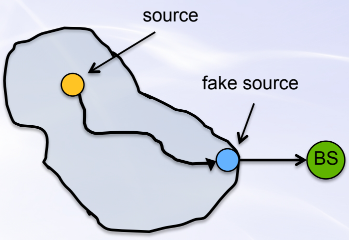


NICS

83

URW: Clouds of fake messages

- The walking phase in [Mahm11,12] is hidden within a cloud of fake messages with irregular shape
- Clouds are activated by real packets traveling to fake sources
- Fake sources are chosen during setup using discovery messages
  - The node chooses a subset of nodes at distance h
  - The response includes the path and a random number R used for generating (chains of) pseudonyms between neighbors in the route



$$id_{AB}^{(1)}, id_{AB}^{(2)}, id_{AB}^{(3)}, \ldots., id_{AB}^{(n)},$$

Where: $id_{AB}^{(i)} = H(K_{AB}, id_{AB}^{(i-1)})$ and $id_{AB}^{(1)} = H(K_{AB}, R)$.

NICS

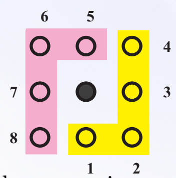
84

- Undirected random walks (URW)
- Directed random walks (DRW)
- Network Loop methods (NL)

NICS

Directed Random Walks (DRW)

- Directed Random Walks (DRW) were introduced to guide the walking phase and thereby circumvent some of the problems derived from using pure random walks
  - Similar consecutive paths
  - Phantom sources close to the real source node
- The following solutions aim to enhance the basic mechanism devised by Phantom Routing
  - Grouping neighbours in two groups
    - Closer
    - Further

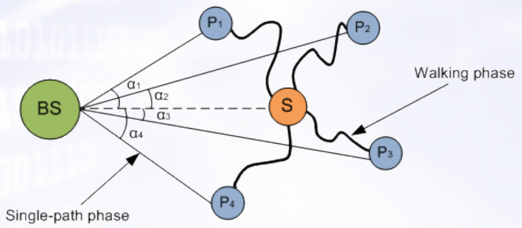


neighbor grouping method

NICS

DRW: Phantom Routing with Location Angle

- PRLA [Weip08] introduces inclination angles to direct random walks
  - Increasing the length of a random walk is useless if the phantom source is not placed in a secure place
  - An attacker placed in the shortest path from BS to S will have a better chance to success if angles of arrival are less pronounced



- Phantom sources with a larger inclination angle are prioritized

NICS 87

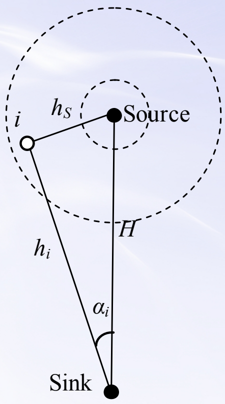
DRW: Phantom Routing with Location Angle

- The source node broadcasts a message for *TTL* hops for nodes in the vicinity to calculate their own inclination angle (by law of cosines):

$$\alpha_i = \arccos \frac{H^2 + h_i^2 - h_s^2}{2Hh_i}$$

  - These values are shared between neighbours to choose the next hop in the path

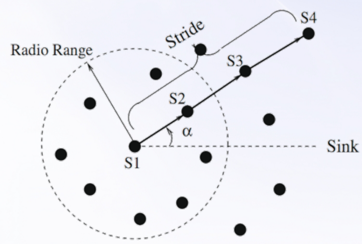
- This process improves safety period but increases communication overhead



NICS 88

DRW: Weighted Random Stride

- WRS [Wang09] is similar to PRLA in the sense that it chooses the next hop in the communication path based on the angle
  - When a sensor transmits data to the BS it first picks a random angle and a stride
    - The stride defines the number of hops associated to the forwarding angle
    - When the stride is finished, the recipient chooses a new forwarding angle and starts a new stride
  - Nodes are designed to pick larger forwarding angles with higher probability

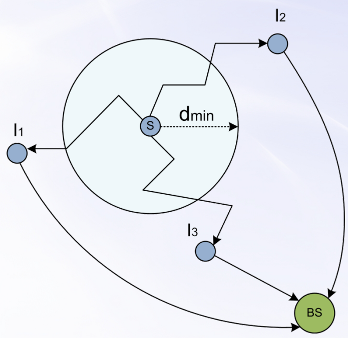


NICS

89

DRW: Routing via Random Intermediate Nodes

- The strategy adopted by [LiRe09a] is to choose a random intermediate node (RRIN) in such a way that they don't stay close to the source
- A node at (x_0, y_0) first chooses a random distance d_{rand} such that $d_{rand} \geq d_{min}$
- Then chooses a random relative location (x_d, y_d), located outside the range of d_{min}, from where the packet will be routed to the BS
  - The node closest to this position will be used as intermediate node

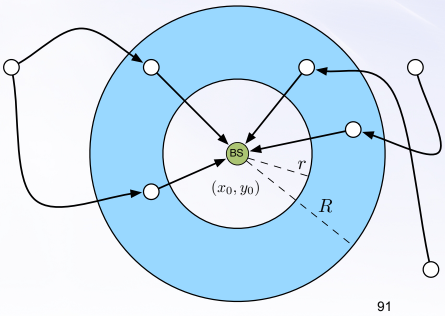


NICS

90

DRW: STaR routing

- The STaR [Ligh10] aims to reduce the cost associated to the selection of pure random intermediate nodes in the field
  - Results in large communication paths
- Instead, RRIN nodes are uniformly and randomly chosen within a toroidal region around the base station
  - $(x_i, y_i) = (x_0 + d\cos\theta,\ y_0 + d\sin\theta)$
- The RRIN finally forwards the packet to the sink using single-path routing

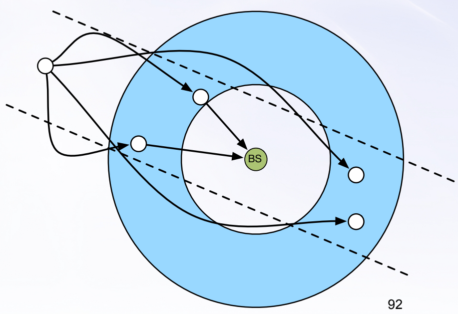


NICS

91

DRW: STaR routing

- The design is intended to give the illusion that the source node is sending messages from all possible directions
- By limiting the area from where random nodes are selected STaR reduces the energy consumption compared to RRIN
  - However, it is not clear whether it is efficient to reach nodes behind the sink
  - Also, this scheme presents the problem described by PRLA wrt to the selection of intermediate nodes near the shortest path between the source and the sink



NICS

92

- Undirected random walks (URW)
- Directed random walks (DRW)
- Network Loop methods (NL)

NICS

93

NL: Network Loops

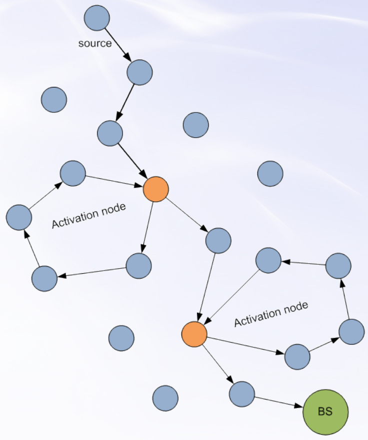
- Network loops consist of a sequence of nodes that transmit fake messages that cycle along the loop
- The goal is to mislead the adversary from the real path of messages and thereby increasing the safety period (i.e., the time it takes for the adversary to reach the source)

NICS

94

NL: Cyclic Entrapment Method

- CEM [Ouya06] aims to trap adversaries into network loops and keep the adversary away from source as much time as possible
- After deployment, each node decides whether it will generate a loop with probability p
  - The node selects two neighbours $\mathcal{A}$, $\mathcal{B}$
  - Sends the packet to $\mathcal{A}$ and after h hops it is delivered to $\mathcal{B}$
  - All intermediaries become loop members
- Loops are activated when a real packet being routed from the source to the sink encounters a loop member (activation node)

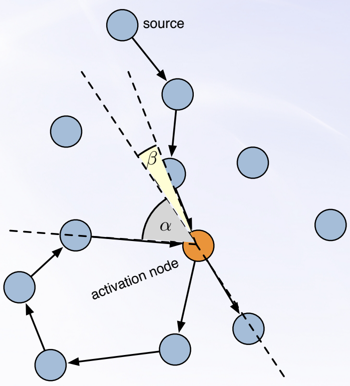


NICS

95

NL: Cyclic Entrapment Method

- Benefits of CEM
  - Real traffic routed normally, no extra delays
  - Fake traffic (i.e., network loops) are deactivated as soon as the loops stops receiving real traffic
- Protection level
  - Depends on the number of active loops (energy trade-off)
  - The attacker is forced to choose which path to follow "randomly" from all the packets observed at the activation node
  - However, he might deduce the right direction by checking shortest-path deviation

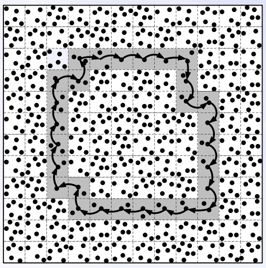


NICS

96

NL: Network Mixing Ring

- The NMR [LiRe09b] builds a ring around the base station which receives real traffic that is mixed with fake traffic before it is finally relayed to its destination
- The communications within the ring have the following features
  - Messages always flow in the clockwise direction
  - Only a few nodes in the ring generate traffic (vehicle messages)
  - Vehicle messages transport several data units, which are all initially fake
  - Fake data units can be replaced with real messages
  - Real message are relayed several hops before exiting the ring
  - Vehicle messages are re-encrypted at every hop

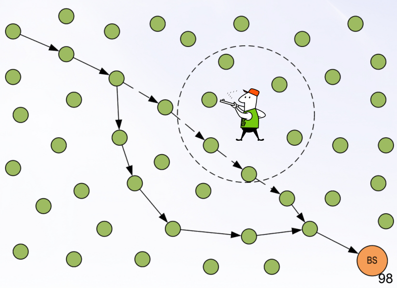


- Real traffic is relayed for a random number of hops to prevent the adversary from learning the entry point to the ring

NICS

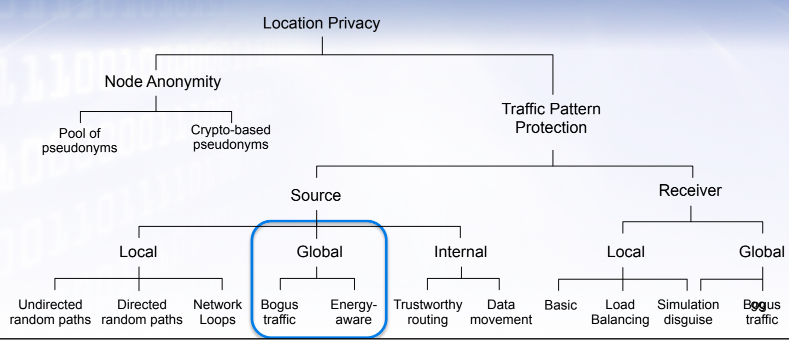
Context-Aware Location Privacy

- Previous solutions are too resource consuming because they are active 24/7
  - We trigger the protection mechanism only in the presence of the adversary
- CALP [Rios11b] benefits from sensors' context-awareness to anticipate the adversary movements
  - Minimize the number of captured packets
  - Minimize the energy consumption



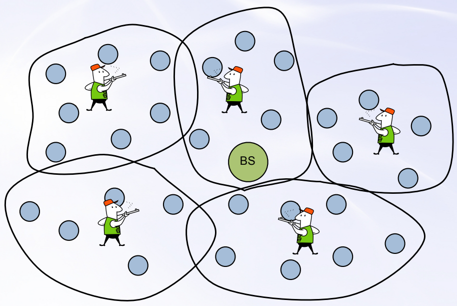
NICS

SLP: Global Adversaries



Global Adversary

- Global attackers are able to monitor the transmission rate of every node in the network
- A global view of the network is usually obtained by several colluding adversaries
  - This can be achieved by deploying an adversarial network covering the sensor field
- Routing-based techniques are known to be ineffective against attackers with a complete view

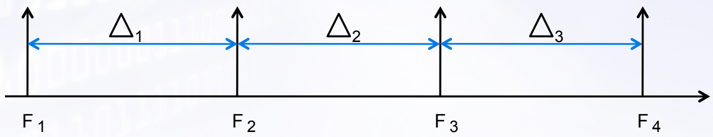


NICS

100

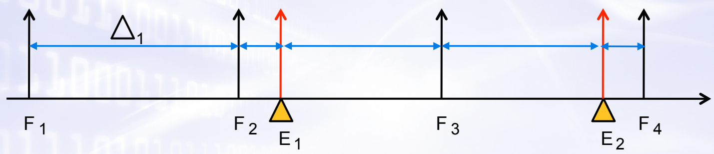
Fake Message Transmission

- A global attacker can easily spot the source of messages because sensor nodes only transmit in the presence of real events
- The idea is to make every node to transmit fake messages (F_x) in order to hide the presence of real events within fake transmissions [Meht07]
  - Make the traffic pattern independent of the presence of events

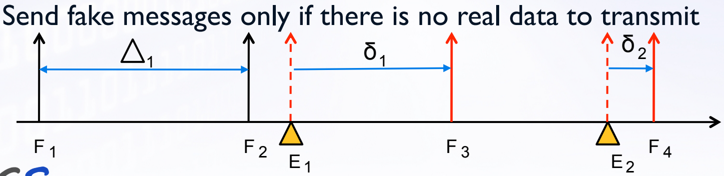


Fake Message Transmission

- Sending fake messages at a constant rate cannot hide the source because the occurrence of a real event message (E_x) will change the fake message transmission pattern

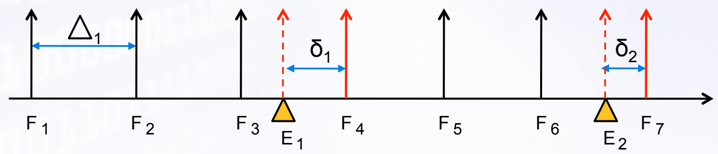


- Periodic Collection [Meht07,12]: Real messages must be delayed (δ) in order to follow the same distribution as fake messages
  - Send fake messages only if there is no real data to transmit



Perfect Event Unobservability

- This method provides the best level of protection (*perfect event source unobservability*), however it might introduce an abusive delivery delay
- Intuitively, the delay can be reduced by reducing the fake inter-transmission times
  - Trade-off between energy consumption and delivery time
    - Large Δ to ensure the durability of the network
    - Low Δ to meet the latency requirements of the application



NICS

103

Problem to solve

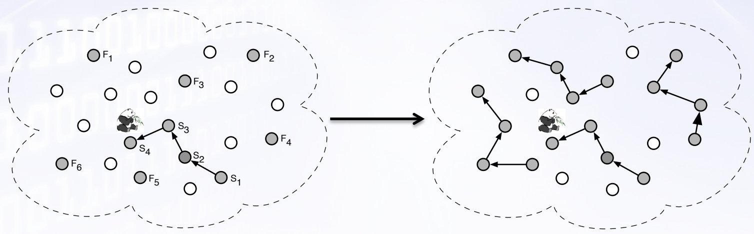
- The problem to be solved is:
  - To provide source location privacy without introducing an excessive delay in nodes transmissions, while preserving nodes batteries
- Some solutions to the problem
  - Source simulation
  - Bogus traffic filtering
  - Statistical approaches

NICS

104

Source Simulation

- To reduce the energy consumption, [Meht07,12] propose to reduce the number of potential sources by creating multiple candidate traces
- Modelling the behaviour of real objects is quite challenging
  - The attacker would be able to easily distinguish fake from real objects if objects are inaccurately simulated



NICS

105

Source Simulation

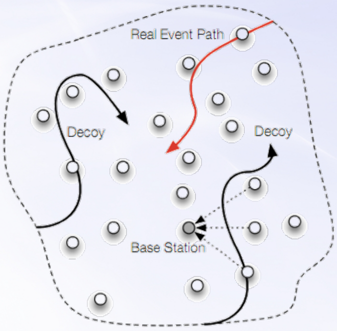
- Mehta *et al.* propose a source simulation protocol as follows:
  - During deployment, a set of L nodes are preloaded with a different token
  - After deployment, token nodes trigger the generation of traffic as if a real event was detected
  - In the next round, the token node passes the token to one of its neighbours (including itself) depending on the behaviour of real objects
  - The behaviour is application-dependent
- The size of the set L determines the privacy level and also the energy consumption

NICS

106

Source Simulation

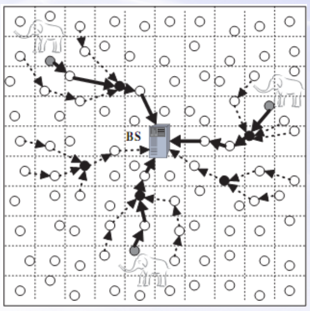
- Unobservable Handoff Trajectory (UHT) [Orto11] aims to protect events originating at the perimeter of the network and eventually expiring inside
  - E.g., transportation of goods
- It consists of a decentralized and self-adaptive scheme that generates fake (mobile) events with the same probability distribution of real events
  - Real events follow a Poisson of ratio l
  - Fake events are generated with rate $k - l$
  - The overall distribution follows a Poisson of ratio k



NICS

Bogus Traffic Filtering

- A set of sensor nodes work as proxies to collect and filter out fake traffic [Yang08]
  - Cells are sending (real or fake) messages at a given rate (i.e., Periodic Collection)
- Upon the reception of traffic a proxy operates as follows
  - Bogus traffic is discarded
  - Real traffic is temporarily buffered and reencrypted

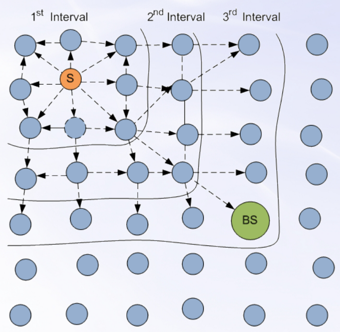


- In case there are no real events available, a proxy sends encrypted dummy messages
  - The attacker cannot learn if the message is real or bogus

NICS

Benefit from existing traffic

- The naïve solution to protect from local adversaries in [Shao09a] is very similar to Period Collection
  - Event data is hidden within beacons, which are periodically sent, thus no extra communications are necessary
  - The main downside is that delivery time increments drastically with the distance from the source to the BS
    - Beaconing intervals are generally large (up to 786 seconds)

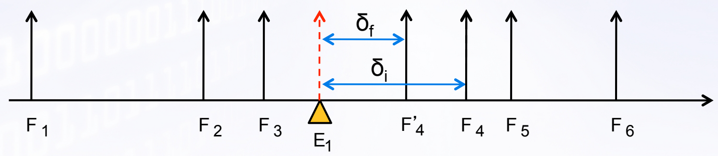


NICS

109

Statistical Approaches

- [Shao08] propose to relax the requirement of perfect event unobservability (Periodic Collection) to statistically strong anonymity to reduce the latency of real events notification
- Given an initial message transmission distribution (F_i), upon the occurrence of a real event (E_I), it can be sent before the next scheduled transmission (F_4)
  - The parameters of the message distribution (e.g., μ, σ) must not be altered



NICS

110

Statistically Strong Source Unobservability

- The Anderson-Darling (goodness of fit) test is used to find an appropriate inter-message delay (*imd*)

Test Statistic: $A^2 = -n - S$, where

$$S = \sum_{i=1}^{N} \frac{2i-1}{n} [\log F(X_i) + \log(1 - F(X_{n+1-i}))].$$

Here F is the CDF of interest, n is the sample size, and X_i denotes the ith datum;

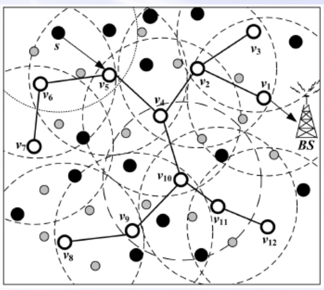
- The search process tries to find the shortest delay that passes the test starting at 0 and gradually increasing this value by a small random number
- As real messages are re-scheduled a.s.a.p., the presence of bursts of events may skew the mean of the distribution
  - The *imd* for a real message is, on average, shorter than the mean
  - This is solved by a mean recovery mechanism, which delays subsequent transmissions

NICS

111

Minimum Connected Dominating Set

- Global eavesdropping is usually achieved by means of an adversarial sensor network deployed to monitor all the transmissions of the network
  - The adversary cannot exactly determine the transmission rate of every particular node
  - Each adversarial node only knows the number of packets sent in its hearing range



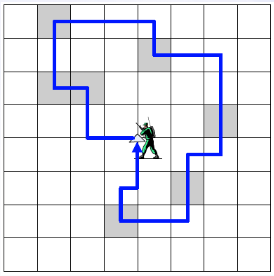
- Therefore, not all sensor nodes need to be active sources of fake traffic [Proa12]
  - Only a subset of nodes act as fake sources
  - Transmissions from the rest of nodes must be controlled
- The subset of fake sources must be of minimum size to reduce the amount of fake traffic

NICS

112

"Active" Global Attack

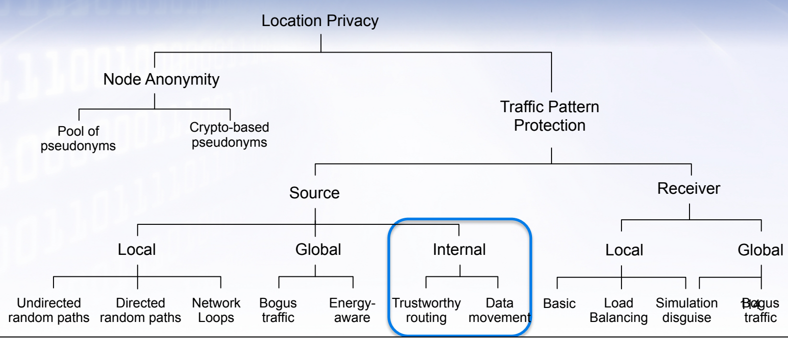
- Previous works consider a passive global attacker in the sense that he doesn't check in the field whether his observations lead to an actual source
- [Yang09] consider a global attacker that upon the detection of suspicious nodes devises an optimal route to visit these spots
- The suspicion level of each cell is determined through traffic analysis
  - The attacker defines a suspicion threshold to determine which cells to visit and in what order
    - Factorial time complexity on the number of suspicious cells $\mathcal{O}(s*s!)$



NICS

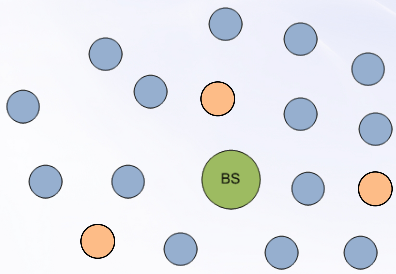
113

SLP: Internal Adversaries



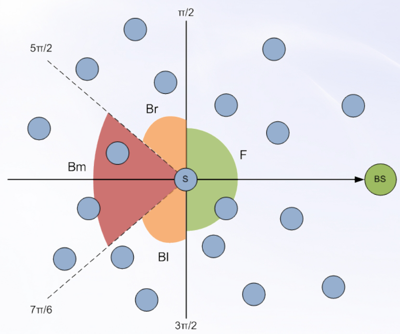
Internal Adversaries

- Active attackers are capable of capturing and compromising several sensor nodes and use them to find the source of event messages
- Internal adversaries are ordinary nodes which are under the control of the adversary
- An internal adversary has access to the crypto material contained in the node and thus it is able to analyse the data contents of the packets traversing it



Trust-based Routing

- [Shai08] propose a trust-based routing algorithm to prevent potentially malicious nodes to forward event data
  - A node calculates a trust value for each neighbour based on the successful interactions with them
  - Each neighbour is classified as trusted or untrusted based
- Additionally, each node classifies its neighbours based on their distance to the BS
  - Forward (F)
  - Backward
    - Right (Br)
    - Left (Bl)
    - Middle (Bm)



Trust-based Routing

- The forwarding process is as follows
  - First, the node picks a random trusted node from the *F* list
  - If no trusted nodes exist it select a random trusted node from *Br* U *Bl*
  - If no trusted nodes exist it chooses a random trusted node from *Bm*
  - If no trusted nodes exist, the packet is dropped
- The identity of the source is protected by replacing the identity at every hop
  - Any intermediate malicious node doesn't know whether the received identity is the real source
- The payload contains the identity of the real source encrypted with the public key of the BS

$$payload = [E_{K_{BS}}(IDx \| rand), E_{K_{XBS}}(data)]$$

NICS

117

Packet alteration schemes

- [Pong11] present SPENA where packets are modified at several random en-route nodes to prevent the association of a packet to the source

DstID	SrcID	Obfuscating Partial Hash (OPH)	Rehash Seed	Payload Length	Payload | SrcID	Filler

- An intermediate node modifies a packet based on the application of some functions to a packet field (i.e., the rehash seed)
  - The packet is modified if $f_p(F_j(seed)) = 1$, where f_p is a mapping function (returns 1 with probability p and 0 with probability 1-p) and F_j is a hash function
- The base station must be able to verify the information and connect it to the source after all modifications

NICS

118

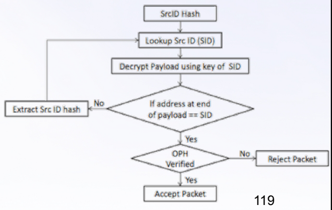
Packet alteration schemes

- The source identifier for every new packet is an element of a hash chain (h_i^m) used in reverse order

DstID	SrcID	Obfuscating Partial Hash (OPH)	Rehash Seed	Payload Length	Payload | SrcID	Filler

- The modifications performed by an intermediate node *j* are
  - The SrcID is replaced by a value of node's j own hash chain (h_j^k)
  - The OPH_i = $[R(h_i^{m+l} \mid Payload)]_{Ki}$ is rehashed and later re-encrypted with its own key (Kj) shared with the base station (i.e., $OPH_j = [R(OPH_i)]_{Kj}$)
  - The $Payload_i$ is replaced by $Payload_j = [R(Payload_i \mid SrcID)]_{Kj}$

- The verification process at destination
  - The BS needs to keep the hash chains of all nodes to find the SrcID and corresponding key
  - Recursively decrypts the payload until it finds the true source
  - Finally, it checks the validity of the OPH



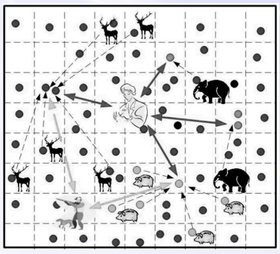
NICS

119

Data movement and anonymous communications

- [Shao09b] concentrate on the problem of source-location privacy in the node presence of node compromise in data-centric sensor networks (DCSN)

- In DCSNs data of different event types are stored at different locations to provide a more efficient access to the data
  - There is not a persistent BS, instead, mobile sinks may collect the stored data on demand based on a publicly known mapping function



- There are two types of sensor nodes in DCSNs
  - *Sensing nodes*: collect and forward information about events of interest
  - *Storage nodes*: receive data of a particular and respond to mobile sink queries

NICS

120

Data movement and anonymous communications

- *p*DCS is designed to prevent an attacker from obtaining event data of his interest. Specifically, it focuses on the prevention of
  - *Node compromise*: retrieval of any the data stored in the node
  - *Mapping attacks:* identify the relation between storage and sensing nodes
- The proposed scheme is based on the use of a secure mapping function and the storage of encrypted data in a remote location
  - Sensing nodes use a mapping scheme based on keyed hash functions to prevent that an attacker determines the location of previous sensed data
    - Future data storage is also protected by a key revocation mechanism
  - Storage nodes are protected because their contents are encrypted with a key that is not present in the node (i.e., the key of the sensing node)
- The flow of data towards storage cells must also be protected by means of some of any anti-traffic analysis technique

NICS

121

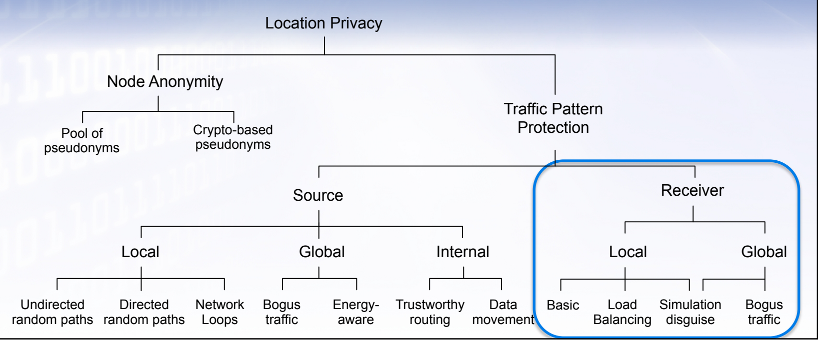
Data movement and anonymous communications

- Several mapping functions are defined
  - *Group-key-based mapping*: all nodes store the same type of event E in the same cell (Lr, Lc) based on a group-wide key K
    - » Lr = H(0|K|E) mod Nr; Lr = H(0|K|E) mod Nc
  - *Time-based mapping*: introduce a group-wide key K_T which is updated periodically after a time period T ($K_T = H(K_{T-1})$)
    - » Lr = H(0|K_T|E|T) mod Nr; Lc = H(1|K_T|E|T) mod Nc
  - *Cell-based mapping*: instead of a network-wide key, each cell (Li, Lj) has its own key K_{ij}, which is also regularly updated
    - » Lr = H(0|i|j|K_{ij}|E|T) mod Nr; Lc = H(1|i|j|K_{ij}|E|T) mod Nc
- These functions are defined in order of increasing privacy because a single node compromise reveals less information (i.e., the location of storage nodes for a set of sensing nodes), which is valid for a shorter time period

NICS

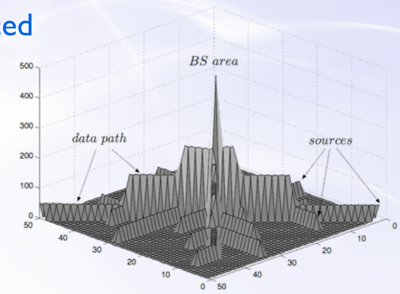
122

Receiver Location Privacy (RLP)



Receiver-Location Privacy

- Refers to the protection of the destination of messages
- The traffic pattern is very pronounced
  - Direction: communications flow in relatively fixed paths
  - Rate: the volume of traffic is higher in the proximities of the base station



- The base station is important for both physical and strategic issues

NICS

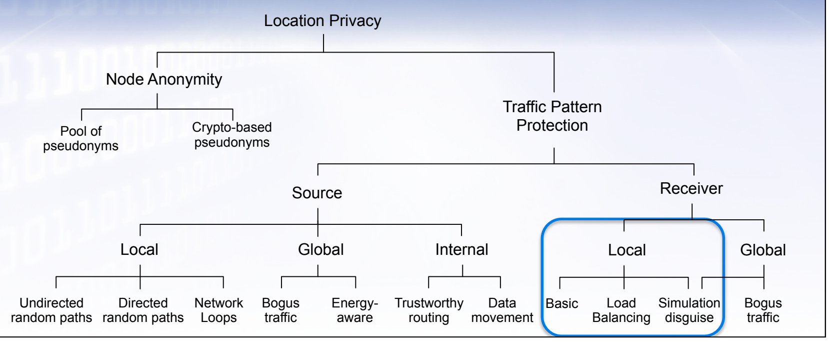
124

Receiver-Location Privacy

- Intuitively, the solution is to homogenize the traffic load on the network
  - Messages must not always follow the shortest path to the destination
  - Every single node should forward a similar number of messages
- Flooding-based protocols provide the maximum homogeneity but at the maximum cost
  - All input messages are forwarded to all neighbours but the sender
- Solutions are also dependent on the power of the adversary

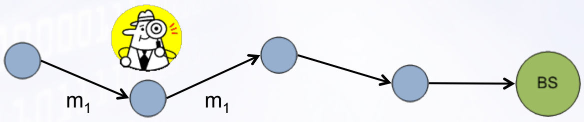
NICS

RLP: Local Adversaries



Basic Countermeasures

- Local attackers are typically placed at a random position in the network and perform different types of attacks [Deng04]
  - Content analysis attacks
  - Time-correlation attacks
  - Rate monitoring attacks
- Content analysis: the attacker can link an incoming packet to an outgoing packet in the same node
  - In sensor networks using shortest-path routing allows the determination of the direction of the communication

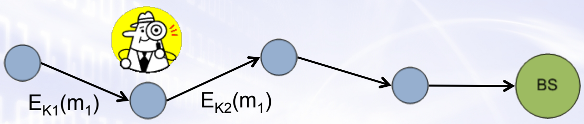


NICS

127

Basic Countermeasures

- Packets indistinguishability prevent packet correlation
  - Apply re-encryption and padding to the messages at each hop



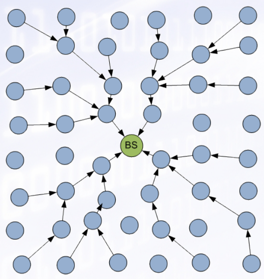
- However, the attacker can also monitor the packet sending times of nodes (time correlation attacks)
  - Apply random delays to packets on their way to the sink
- The attacker might also find the sink by moving towards nodes with a higher transmission rate (rate monitoring attack)
  - Create a uniform sending rate by accepting packets from further nodes only its own packet has been forwarded
  - Otherwise continue to send the same packet or inject dummy traffic if the node has no packet to send

NICS

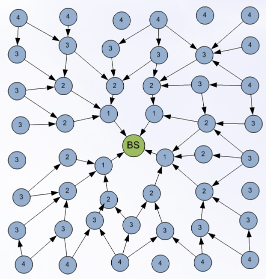
128

Load Balancing Techniques

- There are some limitations to the basic countermeasures that can reduced with traffic-load balancing techniques [Deng06]:
  - Multi-parent routing (MPR): nodes forwards each packet to a random node closer to the base station (parent) balancing the amount of traffic between the different parents, making it more difficult for the adversary to infer the parent-child hierarchy



Single -path routing



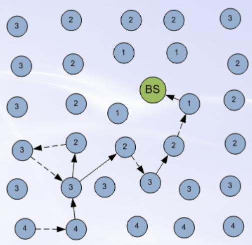
Multi-parent routing

NICS

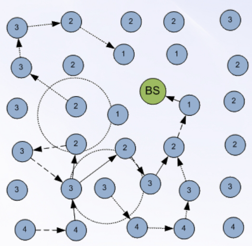
129

Multi-parent routing

- Multi-parent routing (MPR) can be further improved with the addition of
  - Random walks (RW): nodes decide with probability *Pr* whether to send the packet to a random parent or to start a random walk phase with probability 1-*Pr*
    - This addition is intended to mitigate rate monitoring attacks
    - It is still vulnerable to time correlation attacks
  - Fractal propagation (FP): nodes hearing packets in their vicinity inject additional fake messages with a certain probability
    - This mechanism helps to reduce the effect of time correlation attacks



MPR + RW



MPR + RW + FP

NICS

130

Fractal Propagation

- The main problem with fractal propagation (FP) is that nodes in the proximities of the BS generate more traffic
  - The probability of generating fake traffic is the same for all nodes
- Differential FP (DFP) addresses the previous problem by making nodes adapt their probability of generating fake traffic depending on the number of packets they forward
  - This not only reduces the energy consumption and the number of collisions next to the base station but also balances the traffic load more evenly

NICS

131

Simulation and Disguise

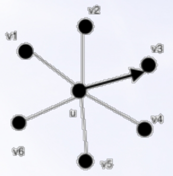
- These solutions attempt to emulate or disguise the presence of the base station at different locations in the field
- Simulation techniques are mainly focused on the creation of *hot spots*, which are areas with a high volume of fake traffic. Several similar approaches have been devised
  - Differential Enforced FP (DEFP) [Deng06]
  - Maelstroms [Chang11]
  - Pseudo-base stations [Biswas08]
- The main challenge is how to create hotspots that are evenly distributed in the network with a minimum overhead

NICS

132

Simulation and Disguise

- Differential Enforced Fractal Propagation (DEFP) [Deng06] is an extension of DFP that generates hotspots in a distributed and dynamic way
- The idea is to make nodes to send fake traffic in the same direction with a higher probability
  - Nodes keep track of the number of fake packets forwarded to its neighbours
  - New fake traffic is more likely to be sent nodes who received more fake traffic before



Ticket table of node u

ID	Tickets	forward probability
v1	1	1/10
v2	1	1/10
v3	5	1/2
v4	1	1/10
v5	1	1/10
v6	1	1/10

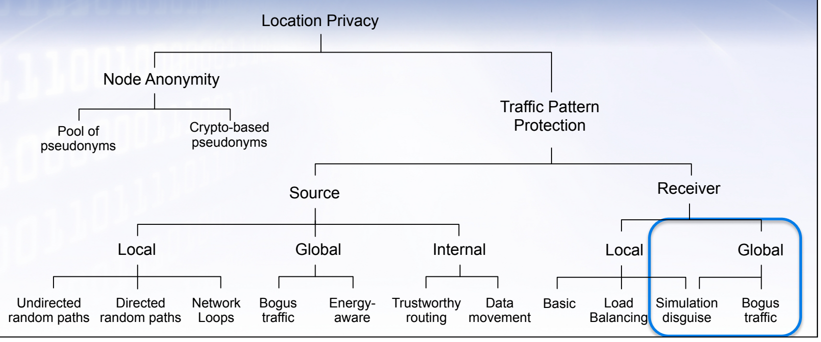
(b) After u forwards a fake packet to node v_3.

- The location of hotspots can be changed on demand by resetting the forwarding probabilities

NICS

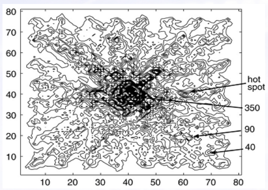
133

RLP: Global Adversaries

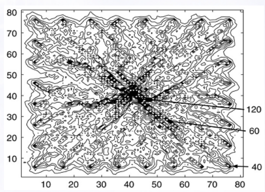


Global Adversaries

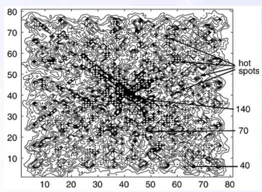
- A global adversary has knowledge about the transmission rate of every sensor node
  - An adversary with real-time analysing capabilities can defeat most of the previous protection mechanisms
- However, if the adversary can only retrieve a snapshot of the amount of traffic generated during a timeslot, previous techniques might provide some means of protection



FP



DFP



DEFP

NICS

135

Global Adversaries

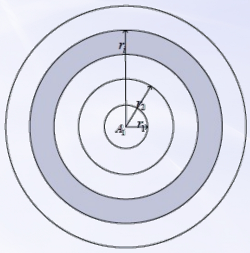
- Again, the injection of fake traffic is one of the main solutions to protect against global adversaries
- Controlling the transmission rate of nodes
  - Use of buffering techniques in the vicinity of the base station
  - Fake packet generation in nodes far from the base station
- Other solutions:
  - Making the base station mimic the behaviour of sensor nodes
  - Simulating the presence of several base stations
  - Moving the base station to a different location

NICS

136

Homogenizing the number of transmissions

- [Ying11a] propose to make all nodes transmit, on average. the same number of packets regardless of their distance to the base station
  - Prevent rate-monitoring by injecting fake traffic on a regular basis
- Each node generates *fm*(i) fake packets and discard those received from its neighbours
  - The rate depends on its distance to the sink



$$fm(i) = TPN_1 - TPN_i = \frac{2ih^2 - 2h^2 + (i-1)^2}{2i-1}$$

  - TPNi is the ratio between all the traffic generated by all rings ≥ i and the number of nodes at ring i
  - h is the maximum distance from a node to the sink
- Nodes are assumed to have a similar transmission rate of real message but what if there are burst of events to transmit?

NICS

137

Homogenizing the number of transmissions

- A similar approach is presented In [Ying11b], where they calculate the transmission rate of nodes based on the number of children nodes a neighbour of the sink has
- The total amount of traffic transmitted by any node is calculated as

$$rate_{Total} = fm(i) + \rho(Ch(i) + 1)$$

  - fm(i) = rate of fake messages
  - Ch(i) = number of children the node has
  - ρ = average rate of real messages
- The fake traffic rate is such that all nodes transmit transmit as much traffic as its one-hop neighbour (i.e., *sink_neigh(i)*)

$$fm(i) = \rho(Ch(sink\_neigh(i) - Ch(i) - 1)$$
$$= Ch(sink\_neigh(i))$$

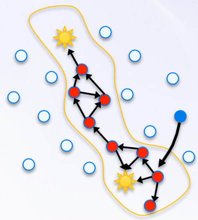
- The authors argue that the lifetime of the network is not affected because the batteries of all nodes are exhausted at the same time

NICS

138

Controlled Flooding

- Previously, we mentioned that flooding-based protocols are the best protection mechanism but are also very costly
- Backbone flooding [Mehta12] reduces the communication cost associated with flooding-based protocols by limiting the scope of the flooding
  - Packets are flooded only among backbone members
- The backbone is created such that
  - The backbone consists of enough sensor nodes to achieve the desired level of privacy
  - Each of the sinks are within the range of at least one backbone member

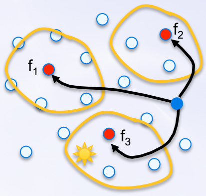


NICS

139

Simulation and Disguise

- Similar to source simulation, which was proposed to counter a global adversary, [Mehta12] also proposes sink simulation
  - Several fake sinks are created to confuse the adversary
- During deployment several sinks are *manually* placed in the field and a subset of sensors are chosen to behave as fake sinks
  - Each real sink must have a fake sink within its communication range
  - All network traffic is addressed to fake sinks, which on reception locally broadcast the message
  - There should be more fake sinks than real sinks



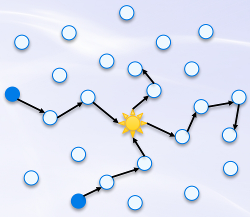
- A source node sends event data to all fake sinks, which perform a single-hop broadcast of messages
  - The adversary might think that a real sink could be nearby but he "only" needs to check the vicinity of *k* fake sinks

NICS

140

Simulation and Disguise

- The idea behind BAR (BS Anonymity via Re-transmission) [Acha10] is that the BS forwards received packets selectively to random nodes nearby
- After receiving a packet the BS decides whether to send the packet on a random walk for a given number of hops *M*



- The value of *M* is dynamically adjusted based on the level of threat perceived by the BS
  - A higher hop count results in a better distribution of packet transmissions in the network
- The attacker is assumed to control solely the transmission rates but not the direction of packets
  - Time correlation attacks could help in deducing the location of the BS

NICS

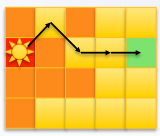
141

Simulation and Disguise

- [Acha10] propose RIA (Relocation for Increased Anonymity), which consists of moving the base station to a safer location
- The new location is calculated based on the impact over network performance and the protection level of the BS
  - The network is divided into cells and the BS knows the transmission rate of each cell and its density (i.e., the number of nodes in the cell)
  - The BS can calculate a score for each cell and move to the cell with the highest value

$$score_i = density_i / threat_i$$

- When moving the BS to its new location using the shortest path saves energy but may be dangerous
  - Instead, the BS follows the least risky path to reach the final location

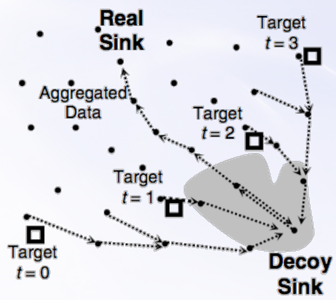


NICS

142

Simulation and Disguise

- The Decoy Sink Protocol [Conner06] uses indirection and data aggregation to reduce the amount of traffic received by the real sink
- Sensors nodes send their data to a decoy sink and on their way the data are aggregated and finally the decoy sink forward the aggregation to the real sink
- The protocol is extended to use several randomly deployed decoy sinks
  - The attacker can discard to look for the real sink in areas where the traffic rate is high
  - Several decoy nodes result in a better balance of network traffic
  - All sensors send data to the same decoy sink during the same time period
- The attacker model only considers rate monitoring attacks



NICS

143

FINAL REMARKS

144

Final Remarks

- Privacy in WSNs poses new challenges because of the nature of the networks and the lack of protection provided by traditional security mechanisms
- Location privacy solutions are mainly based on
  - Routing-based protocols to counter local adversaries
  - Fake message transmissions to provide event unobservability in the presence of global adversaries
  - Little work has been done against internal adversaries

NICS

145

Final Remarks

- There is always a cost associated to the application of privacy preserving techniques which must be carefully taken into consideration when dealing with highly resource-constrained devices
- New scenarios, adversarial models and solutions are expected to appear with the full integration of WSNs and the Internet

NICS

146

References

- [Aky02] “*Wireless sensor networks: a survey*”, Computer Networks, Volume 38, Issue 4, 2002, Elsevier, pp. 393-422
- [Walt2007] Security in Distributed, Grid, and Pervasive Computing, chapter “*Wireless Sensor Network Security: A Survey*”, Auerbach Pub. 2007, pp. 367-409.
- [West67] Westin, A. F. *Privacy and Freedom*, New York Atheneum, 1967.
- [Raym00] Raymond, J.-F. "Traffic Analysis: Protocols, Attacks, Design Issues, and Open Problems", *in* Federrath, H., ed.,'Proceedings of Designing Privacy Enhancing Technologies: Workshop on Design Issues in Anonymity and Unobservability', Springer-Verlag, LNCS 2009, 2000, pp. 10--29.
- [Pfit10] Pfitzmann, A. and Hansen, M. "A terminology for talking about privacy by data minimization: Anonymity, Unlinkability, Undetectability, Unobservability, Pseudonymity, and Identity Management", v0.34, 2010.
- [Rios12a] Rios, R. and Lopez, J. "(Un)Suitability of Anonymous Communication Systems to WSN," *IEEE Systems Journal, In Press.*
- [Chau81] Chaum, D. "Untraceable Electronic Mail, Return Addresses, and Digital Pseudonyms," *Commun. ACM* (24:2), 1981, pp. 84--88.
- [Reit98] Reiter, M. and Rubin, A. "Crowds: Anonymity for Web Transactions," *ACM transactions on information and system security* (1:1), 1998, pp. 66--92.

NICS

147

References

- [Chau88] Chaum, D. "The Dining Cryptographers Problem: Unconditional Sender and Recipient Untraceability," *Journal of Cryptology* (1), 1988, pp. 65--75.
- [Pai08] Pai et al."*Transactional Confidentiality in Sensor Networks*," IEEE Security & Privacy, Volume 6, Issue 4, 2008, pp. 28-35.
- [Mis06] Misra & Xue, "Efficient anonymity schemes for clustered wireless sensor networks," *International Journal of Sensor Networks (1:1), 2006, pp. 50--63.*
- [Ouya07] Ouyang, et al."Providing Anonymity in Wireless Sensor Networks“ 'Pervasive Services, IEEE International Conference on', 2007, pp. 145-148.
- [Rios11a] Rios, R. and Lopez, J. "Analysis of Location Privacy Solutions in Wireless Sensor Networks," *IET Communications, IET Journals,* Vol 5. Issue 17, pp 2518-2532
- [Oztu04] Ozturk et al. "*Source-location privacy in energy-constrained sensor network routing*”, ACM workshop on Security of ad hoc and sensor networks, 2004, pp. 88--93.
- [Kama05] Kamat et al. "*Enhancing Source-Location Privacy in Sensor Network Routing*”. IEEE International Conference on Distributed Computing Systems, 2005, pp. 599-608.
- [Xi06] Xi and Shi "*Preserving source location privacy in monitoring-based wireless sensor networks*“, Parallel and Distributed Processing Symposium, 2006, pp. 8.

NICS

148

References

- [Wang09] Wang, et al. "Privacy-aware routing in sensor networks", *Comput. Netw., Elsevier* 2009, *53*, 1512-1529
- [Shao09] Shao, et al "Cross-layer Enhanced Source Location Privacy in Sensor Networks"'IEEE Conference on Sensor, Mesh, and Ad Hoc Communications and Networks (SECON '09)', IEEE Communications Society, 2009, pp. 1—9.
- [Mahm12] Mahmoud, M. E. and Shen, X. "Secure and Efficient Source Location Privacy-Preserving Scheme for Wireless Sensor Networks"'Proceedings of the IEEE International Conference on Communications (ICC'12)', IEEE Communications Society, Ottawa, Canada, 2012.
- [Mahm11] Mahmoud, M. and Shen, X. "A Cloud-Based Scheme for Protecting Source-Location Privacy against Hotspot-Locating Attack in Wireless Sensor Networks," *IEEE Transactions on Parallel and Distributed Systems (PP:99), 2011, pp. 1.*
- [Weip08] Wei-ping et al."*A Source-Location Privacy Protocol in WSN Based on Locational Angle*", IEEE International Conference on Communications, 2008, pp. 1630-1634.
- [LiRe09a] Li & Ren, "*Providing Source-Location Privacy in Wireless Sensor Networks*", *WASA '09: Proceedings of the 4th International Conference on Wireless Algorithms, Systems, and Applications, Springer-Verlag,* 2009, 338-347
- [Ligh10] Lightfoot, L., Li, Y. and Ren, J. "Preserving Source-Location Privacy in Wireless Sensor Network Using STaR Routing.GLOBECOM'10', 2010, pp. 1-5.
- [Ouya06] Ouyang et al. "*Entrapping Adversaries for Source Protection in Sensor Networks*" International Symposium on World of Wireless, Mobile and Multimedia Networks, IEEE Computer Society, 2006, pp. 23—34.
- [LiRe09b] Li & Ren, "*Preserving Source-Location Privacy in Wireless Sensor Networks*", SECON'09: Proceedings of the 6th Annual IEEE communications society conference on Sensor, Mesh and Ad Hoc Communications and Networks, IEEE Press, 2009, 493-501

NICS

149

References

- [Rios11b] Rios, R. and Lopez, J. "Exploiting Context-Awareness to Enhance Source-Location Privacy in Wireless Sensor Networks," *The Computer Journal,* (54):10, 2011, pp.1603-1615.
- [Meht07] Mehta et al."*Location Privacy in Sensor Networks Against a Global Eavesdropper*", IEEE International Conference on Network Protocols, 2007, pp. 314-323.
- [Ortolani 2011] Ortolani, S., Conti, M., Crispo, B. and Di Pietro, R. "Events privacy in WSNs: A new model and its application"'IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)', 2011, pp. 1 -9.
- [Shao08] Shao et al."*Towards Statistically Strong Source Anonymity for Sensor Networks*", INFOCOM 2008. The 27th Conference on Computer Communications. IEEE , 2008, pp. 466-474.
- [Yang08] Yang et al."*Towards event source unobservability with minimum network traffic in sensor networks*" WiSec '08: Proceedings of the first ACM conference on Wireless network security', ACM, 2008, pp. 77--88.
- [Alom09] Alomair, et al.,"*Statistical Framework for Source Anonymity in Sensor Network*" (003), Technical report, Network Security Lab (NSL), 2009.
- [Proano12] Proano, A. and Lazos, L. "Hiding contextual information in WSNs", 'IEEE International Symposium on a World of Wireless, Mobile and Multimedia Networks (WoWMoM)', 2012, pp. 1 -6.
- [Ying11a] Ying et al., "*Concealing of the Sink Location in WSNs by Artificially Homogenizing Traffic Intensity*",The First International Workshop on Security in Computers, Networking and Communications (INFOCOM Workshops), 2011, pp. 988 - 993.
- [Ying11b] Ying et al., "*A Protocol for Sink Location Privacy Protection in Wireless Sensor Networks*",IEEE Global Telecommunications Conference (GLOBECOM), IEEE Communications Society, Houston, TX, USA, 2011, pp. 1 -5.

NICS

150

References

- [Yang09] Yang, Y. et al. "An Active Global Attack Model for Sensor Source Location Privacy: Analysis and Countermeasures", *in 'Security and Privacy in Communication Networks', Springer Berlin Heidelberg, 2009, pp. 373-393.*
- [Shai08] Shaikh et al. "*Network Level Privacy for Wireless Sensor Networks*" 'ISIAS '08. Fourth International Conference on Information Assurance and Security.', 2008, pp. 261-266.
- [Pong11] Pongaliur, K. and Xiao, L. "Maintaining source privacy under eavesdropping and node compromise attacks"'Proceedings IEEE INFOCOM 2011', 2011, pp. 1656 -1664.
- [Shao 2009] Shao et al., "pDCS: Security and Privacy Support for Data-Centric Sensor Networks", *Mobile Computing, IEEE Transactions on (8:8), 2009, pp. 1023-1038.*
- [Deng05] Deng et al., "*Countermeasures Against Traffic Analysis Attacks in Wireless Sensor Networks*", SECURECOMM '05, IEEE Computer Society, 2005, pp. 113--126.
- [Deng06] Deng et al., "*Decorrelating wireless sensor network traffic to inhibit traffic analysis attacks*", Pervasive and Mobile Computing, Volume 2, Issue 2, 2006, pp. 159--186.
- [Rios12b] Rios, R., Cuellar, J., Lopez, J., "*Robust Probabilistic Fake Packet Injection for Receiver-Location Privacy in WSN*", ESORICS 2012, pp. 163-180, 2012.
- [Jian07] Jian et al., "*Protecting Receiver-Location Privacy in Wireless Sensor Networks*" INFOCOM 2007. 26th IEEE International Conference on Computer Communications, 2007, pp. 1955-1963.

NICS

References

- [Bisw08] Biswas, et al, "*A Countermeasure against Traffic-Analysis based Base Station Detection in WSN*". [online] http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.98.948
- [Chang11] Chang, S., et al. "Maelstrom: Receiver-Location Preserving in Wireless Sensor Networks", *in Cheng, Y., Eun, D., Qin, Z., Song, M. and Xing, K., ed.,'Wireless Algorithms, Systems, and Applications', Springer Berlin / Heidelberg, 2011, pp. 190-201.*
- [Acha10] Acharya and Younis, "*Increasing base-station anonymity in wireless sensor networks. Ad Hoc Networks*", In Press, Uncorrected Proof, 2010.
- [Conner06] Conner, W., Abdelzaher, T. and Nahrstedt, K. "Using Data Aggregation to Prevent Traffic Analysis in Wireless Sensor Networks", *in Gibbons, P., Abdelzaher, T., Aspnes, J. and Rao, R., ed.,'Distributed Computing in Sensor Systems', Springer Berlin / Heidelberg, 10.1007/11776178_13, 2006, pp. 202-217.*
- [Li09] Li, X.et al., "Enhanced Location Privacy Protection of Base Station in Wireless Sensor Networks," *Mobile Ad-hoc and Sensor Networks, International Conference on (0), 2009, pp. 457-464.*

NICS

Thanks for your attention!

Javier Lopez
Computer Science Department
University of Malaga
Spain

jlm@lcc.uma.es

NICS