Secure location determination and verification in wireless networks

- 1. Lazos, L. and Poovendran, R. 2004. SeRLoc: secure range-independent localization for wireless sensor networks. In Proceedings of the 3rd ACM Workshop on Wireless Security (Philadelphia, PA, USA, October 01 - 01, 2004). WiSe '04.
- 2. S. Ray, R. Ungrangsi, F. D. Pellegrini, A. Trachtenberg, and D. Starobinski. Robust location detection in emergency sensor networks. In Proceedings of IEEE INFOCOM 2003, April 2003.
- 3. Liu, D., Ning, P., and Du, W. K. 2005. Attack-resistant location estimation in sensor networks. In Proceedings of the 4th international Symposium on information Processing in Sensor Networks (IPSN) (Los Angeles, California, April 24 - 27, 2005).
- 4. Ekici, E.; Mcnair, J.; Al-Abri, D., "A Probabilistic Approach to Location Verification in Wireless Sensor Networks," Communications, 2006. ICC '06. IEEE International Conference on , vol.8, no., pp.3485-3490, June 2006.

SeRLoc: Secure Range-Independent Localization for Wireless Sensor Networks.

Dependable Computing Systems Laboratory

DongHoon Shin

19 September 2007

Contents

- Introduction
- Network Model
- Localization Algorithm: SeRLoC
- Security Mechanisms of SeRLoc
- Threat Analysis
- Performance Evaluation
- Conclusion

3

Introduction

- Wireless ad hoc sensor networks (WSN)
  - Operate in the absence of a pre-deployed infrastructure
  - Self-configurable
  - Low cost
  - Rapidly deployed
- Localization for WSN
  - Many of the applications proposed for WSN require knowledge of the origin of the sensing information.
    - e.g., the location of the sensors that detect high stress forces is needed in order to identify a crack in the arch of a bridge
  - Furthermore, location is assumed known in the realization of many network operations.
    - Routing protocols where a family of geographically aided algorithms
    - Security protocols where location information is used to prevent threats

4

Introduction

- WSN in hostile environments
  - An adversary can interrupt the functionality of location-aware applications by exploiting the vulnerabilities of the localization scheme.
- Contributions of this paper
  - Propose SeRLoc, a novel range-independent localization scheme
    - Decentralized, resource-efficient sensor localization
  - Propose security mechanism for SeRLoc
    - Allow each sensor to determine its location even in the presence of well known threats on WSN such as wormhole attack, sybil attack, and sensor compromise
  - Provide simulation studies
    - Comparison with state-of-the-art decentralized range-independent localization schemes

5

Network Model

- Network generation
  - A set of sensor nodes of unknown location and a set of locators are deployed randomly in a specific network region of area A.
    - Locators are specially equipped nodes with known location and orientation.
    - Position of locators can be acquired through GPS receivers.
  - # of sensor nodes: N, # of locators: L
  - Communication range from locator to sensor: R
  - The random deployment of the network nodes is modeled as a spatial homogeneous Poisson point process.
  - Let LH_s be the set of locators heard by a sensor s.
  - The probability that s hears exactly k locators is given by

$$P(|LH_s| = k) = \frac{(\rho_L \pi R^2)^k}{k!} e^{-\rho_L \pi R^2}$$

6

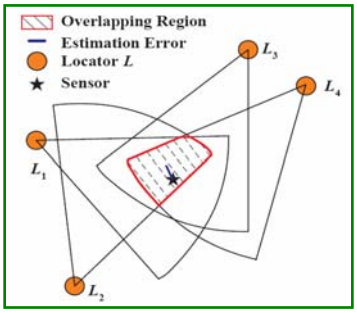
Network Model

- Antenna model
  - Sensors are equipped with omnidirectional antennas.
    - Sensor-to-sensor communication range: r
  - Locators are equipped with sectored antennas with M sectors.
  - Locators transmit with higher power than sensors, hence R>r.
  - Locators can simultaneously transmit in all their antenna sectors.
- Additional assumptions
  - Sensors and locators can be pre-loaded with cryptographic quantities before deployment.
  - Locators are trusted and cannot be compromised by an adversary.

7

Location Determination

- Idea of scheme
  - Each locator transmits different beacons at each antenna sector with each beacon containing
    - Locator's coordinates
    - Angles of the antenna boundary lines, with respect to a common global axis
  - A sensor heard locators computes overlapped regions.
  - Estimated the sensor's location is the center of gravity (CoG) of the overlapping region.
    - Why CoG? → It is the least square error solution given that a sensor can lie with equal probability at any point of the overlapping region.



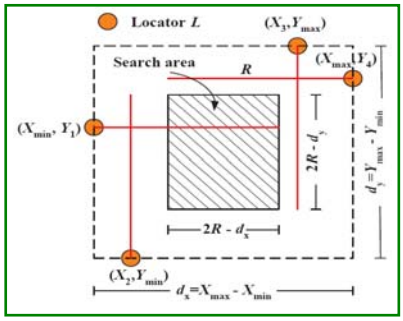
8

Location Determination

- Step1: Locators heard
  - Sensor s collects the beacons from all locators it can hear.

$$LH_s = \{\|s - L_i\| \leq R, \quad i = 1 \ldots |L|\}$$

- Step2: Search area
  - The sensor s computes a search area where it will attempt to locate itself.
  - Initially, s finds the following locator coordinates from LH_s;
    X_{min} = min X_i, X_{max} = max X_i,
    Y_{min} = min Y_i, Y_{max} = max Y_i.
  - The search area A_s is given by
    $A_s = \{(X,Y) | X_{min}\text{-}R \leq X \leq X_{max}\text{-}R,$
    $Y_{min}\text{-}R \leq Y \leq Y_{max}\text{-}R\}$



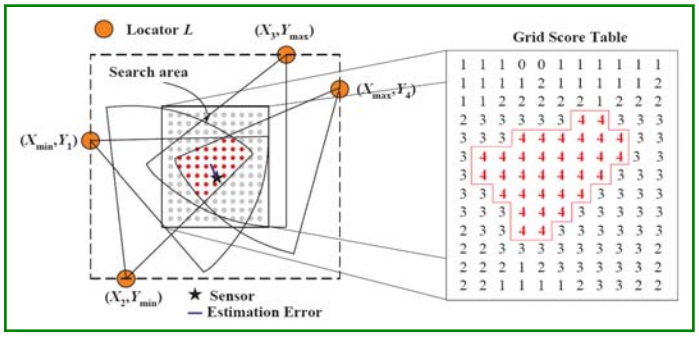
9

Location Determination

- Step3: Overlapping region-Majority vote
  - The sensor s determines the overlapping region of all sectors.
  - It would be expensive for each sensor to attempt to analytically determine the overlapping region, based on the line intersections.
  - A grid scoring system is employed that defines the overlapping region based on majority vote.
  - Grid score table
    - The sensor s places a grid of equally spaced points within the rectangular search area, A_s.
    - The sensor s keeps a score for every grid point in a grid score table to determine the overlapping area.
    - If a point is included in a sector according to a grid sector test described below, the sensor s increments its score by one; otherwise, its score does not change.
    - This process is repeated for all locators heard, LH_s, then finally, the overlapping region is defined by the points with the highest score.

10

Location Determination

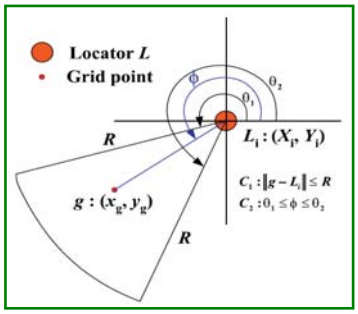


- The resolution of the grid can be increased to reduce the error at the expense of energy consumption due to the increased processing time.

Location Determination

- Grid-sector test
  - Let the coordinates of a grid point g be denoted as (x_g, y_g).
  - Point g is included in a sector of angles $[\Theta_1, \Theta_2]$ originating from locator $L_i : (X_i, Y_i)$ if it satisfies two conditions:
    (C_1): g has to lie within the communication range of L_i.
    (C_2): The angle ϕ of the line connecting L_i and g, has to lie within $[\Theta_1, \Theta_2]$.

$$C_1 : \ \|g - L_i\| \le R,$$
$$C_2 : \ \theta_1 \le \phi \le \theta_2$$



- Note that the sensor does not have to perform any angle-of-arrival (AOA) measurements

Location Determination

- Step4: Location estimation
  - The sensor s determines its location as the centroid of all the grid points that define the overlapping region (highest score in the grid):

$$\tilde{s} : (x_{est}, y_{est}) = \left(\frac{1}{n} \sum_{i=1}^{n} x_{g_i}, \frac{1}{n} \sum_{i=1}^{n} y_{g_i} \right)$$

13

Security Mechanisms of SeRLoc

- Encryption
  - Sensors and locators share a global symmetric key, K_0, pre-loaded before deployment.
    - With K_0, all beacon messages from locators are encrypted.
  - In addition, every sensor shares a symmetric pairwise key, $K(s,L_i)$ with every locator, L_i, also pre-loaded.
- Locator ID authentication
  - The use of a shared symmetric key does not identify the source of the messages that each sensor hears.
    - A malicious sensor may inject false localization information through the shared key, K_0 to impersonate multiple locators.
  - Sensors are required to authenticate the source of the beacons using collision-resistant has functions.

14

Security Mechanisms of SeRLoc

- Pre-requisites
  - Each locator L_i has a unique PW_i.
  - Due to the collision resistance property, it is infeasible for an attacker to find a value PW_j, such that $H(PW_j) = H(PW_j)$, $PW_i \neq PW_j$.
  - The hash sequence is generated using the following equation:

    $$H^0 = PW_i,\ H^k = H(H^{k-1}),\ k = 1, 2, \ldots, n,$$

    with n being a large number and H^0 never revealed to any sensor.
  - Each sensor is pre-loaded with a table containing the Id of each locator and the corresponding has value $H^n(PW_j)$.

- Authentication of locator's ID
  - Assume that a locator L_i wants to transmit its first beacon.
  - Initially, sensors only know the hash value $H^n(PW_j)$.
  - The locator includes $(H^{n-1}(PW_j), j)$ in the beacon with the index j=1 (first hash value published).

15

Security Mechanisms of SeRLoc

- Every sensor heard the beacon authenticate the locator's ID only if

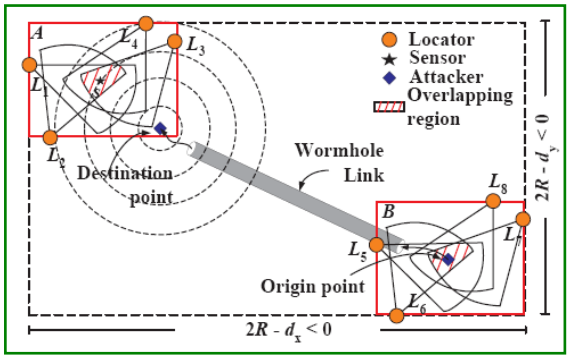
  $$H(H^{n-1}(PW_j)) = H^n(PW_j).$$
- After verification, the sensor places $H^n(PW_j)$ with $H^{n-1}(PW_j)$ in its memory and increases the hash counter by one.
- The beacon of locator L_i has the following format:

  $L_i : \{ (X_i, Y_i) \,||\, (\Theta_1, \Theta_2) \,||\, , (H^{n-j}(PW_j)), j \}_{Ko}$.

16

Wormhole attack against SeRLoc

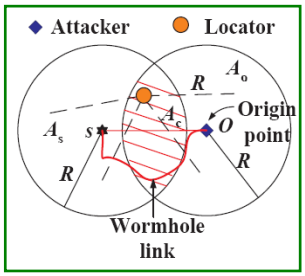
- Wormhole attack against SeRLoc
  - An attacker records beacons at region B, tunnels them via the wormhole link in region A and replays them.
  - The purpose of wormhole attack is to lead the sensor s to believe that it can hear locators {L_1~L_8}.



Wormhole attack against SeRLoc

- Detecting wormholes
  1. Sector uniqueness property
  2. Communication range violation property

- Sector uniqueness property
  - It is infeasible for a sensor to hear two sectors of a single locator.
  - A_s: area where locators heard to sensor s can reside
    A_o: area where locators heard at the origin of the attack can reside
    $A_c = A_s \cap A_o$

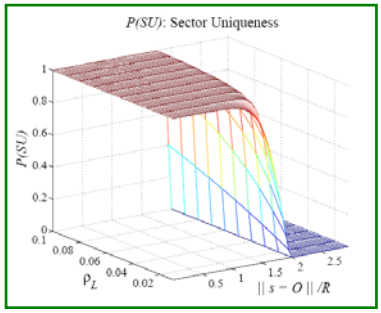


Wormhole attack against SeRLoc

- The detection probability P(SU) is given by

$$P(SU) = P(|LH_{A_c}| \geq 1) = 1 - P(|LH_{A_c}| = 0) = 1 - e^{-\rho_L A_c},$$

where LH_{Ac} denotes the set of locators heard by sensor s that lie inside A_c.

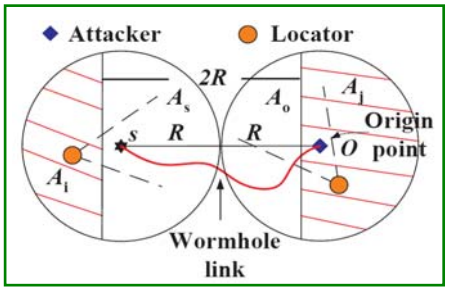


19

Wormhole attack against SeRLoc

- Communication range violation property
  - Every locator directly heard to a sensor s is less than R units away from s, i.e., $\| s - L_i \| < R$, for all L_i in LH_s.
  - Hence, any two locators L_i, L_j in LH_s, heard to s, cannot be more than 2R apart, i.e.,

    $\| L_i - L_j \| \leq \| s - L_i \| + \| s - L_i \| < 2R.$

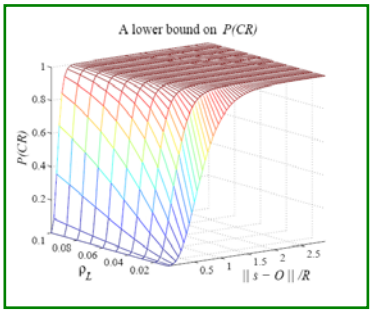


20

Wormhole attack against SeRLoc

- The detection probability P(CR) is given by

$$\begin{aligned} P(CR) &= P(\|L_i - L_j\| > 2R) \\ &\geq P(CR \bigcap (|LH_{A_i}| > 0 \cap |LH_{A_j}| > 0)) \\ &\vdots \\ &= (1 - e^{-\rho_L A_i})(1 - e^{-\rho_L A_j}) \end{aligned}$$



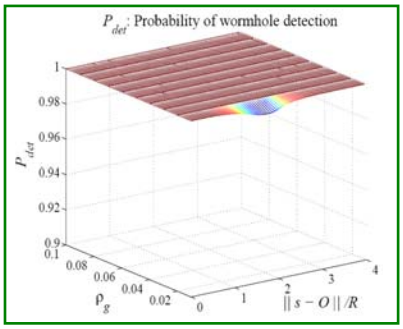
21

Wormhole attack against SeRLoc

- Detection probability
  - By combining two detection techniques, a lower bound on the detection probability P_{dec} of a wormhole attack is computed as:

$$\begin{aligned} P_{det} \geq\ & (1 - e^{-\rho_L A_c}) + (1 - e^{-\rho_L A_i^*})^2 - \\ & (1 - e^{-\rho_L A_c})(1 - e^{-\rho_L A_i^*})^2, \end{aligned}$$

  - The lowest detection prob. is P_{dec} = 99.48%, attained at ρ_L = 0.01.



22

Wormhole attack against SeRLoc

- Location resolution algorithm
  - To resolve the location ambiguity, a sensor under attack executes the *Attach to Closer Locator Algorithm* (ACLA).

Attach to Closer Locator Algorithm (ACLA)

1. $A: \{X \geq X_{min},\ X \leq X_{max},\ Y \geq Y_{min},\ Y \leq Y_{max}\}$.
2. Place a point grid in A and execute the grid score test.
3. Compute $CoG_i\ \forall\ D_i$ with score $\geq th$.
4. $\forall\ D_i$, find $L_i^* \in LH_s$
   $L_i^* = \min \|L_i - CoG_i\|$.
5. $\forall\ L_i^*$, broadcast $\{\eta_i \| q\}_{K_s^{L_i^*}} \| Id_s$.
6. Identify $L_i' \in L^*$ that replies first with the correct nonce.
7. Sensor location: CoG_i' of the region corresponding to L_i'.

To account for the time difference due to transmission time

Sybil Attack & Compromised Sensors

- Threat model
  - In the sybil attack, an attacker impersonates multiple network entities by assuming their identities.
- Sybil attack against SeRLoc
  - Sensors do not rely on other sensors to compute their location. Hence, an attacker has no incentive to assume sensor identities.
  - To impersonate locators, an attacker has to compromise the global key K_o used by locators to transmit beacon.
  - Once K_o has been compromised, the attacker can obtain published values of the hash chains of the locators it hears.
  - Since the sensor always has the latest published values from the locators that it can directly hear, an attacker can only impersonate locators that are not directly heard by the sensor under attack.

Sybil Attack & Compromised Sensors

- Defense against the sybil attack
  - No mechanism is provided to prevent an attacker from impersonating locators except for the ones directly heard to a sensor.
  - However, as long as the pairwise keys between the locators and sensors are not compromised, sensors can still determine the position of them in the presence of a sybil attack.
  - The attacker has no way to decrypt the nonce, encrypted with the pairwise key, or encrypt any kind of reply.

25

Related Work

- Classification of localization scheme
  - Range-dependent scheme vs range-independent based schemes.
- DV-hop
  - Each node discovers the shortest path in number of hops to every other node.
  - Reference points compute the average length of one hop based on the hop count to other reference points, and flood the network with the hop estimate.
  - Nodes use the hop size estimate and the number of hops to compute their distance to at least three reference points and perform triangulation to determine their location.
- Amorphous
  - Employs a similar strategy with the exception of computing average hop size offline through an approximate formula.

26

Related Work

- APIT
  - A sensor relies on neighbor sensor information to determine if it is located inside or outside a virtual triangle defined by three reference point called anchors.
    - For example, a sensor s measures the power to three anchors A, B, C, and also gathers the measurements of all neighboring sensors s_1~s_4.
    - If no neighbor of s is further from to all three anchors simultaneously, s assumes it is outside ΔABC.
    - Otherwise, s assumes it is outside ΔABC.
    - The sensor s repeats the APIT test for all 3-tuples of anchors heard, and estimates its position as the center of gravity of the overlapping region of the triangles for which APIT test was positive.
- Centroid
  - It is outdoor localization scheme, where the reference points broadcast beacons with their coordinates.
  - Nodes estimate their position as the centroid of the locations of all the reference points that they hear.
  - Simple implementation and low communication cost but crude approximation of node location

27

Performance Evaluation

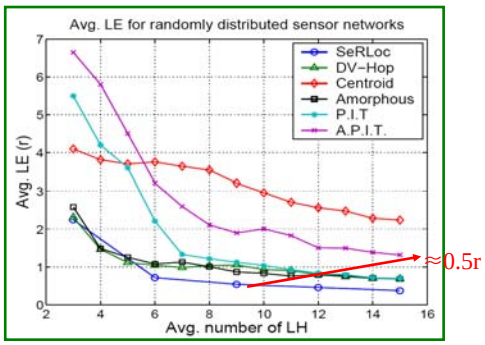
- Simulation setup
  - 5000 sensors within a 100x100 rectangular area
  - Randomly placed locators within the same area
  - Average localization error:

$$\overline{LE} = \frac{1}{|N|}\sum_{i}\frac{\|\tilde{s}_i - s_i\|}{r}$$

28

Performance Evaluation

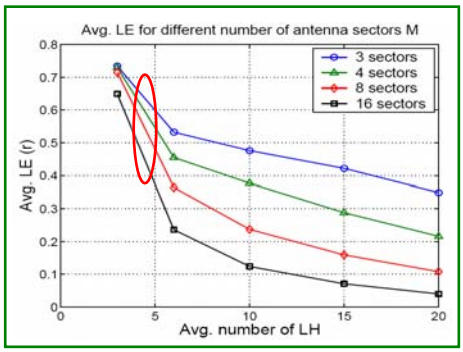
- Localization Error vs. Locators heard
  - For a fair comparison, Avg. LH for SeRLoc is normalized by multiplying avg. LH with # of sectors used.
    - e.g., when (Avg. LH = 9) with SeRLoc using 3 sectors, every sensor hears 3 locators on average
  - SeRLoc is superior to all other algorithms
  - Achieved Avg. LE ≈ 0.5r



29

Performance Evaluation

- Localization Error vs. Antenna Sectors
  - As the Avg. LH increases, the Avg. LE decreases more rapidly for higher number of antenna sectors owing to the smaller overlapping region (because of the narrower antenna sectors)
  - Avg. LE vs. hardware complexity



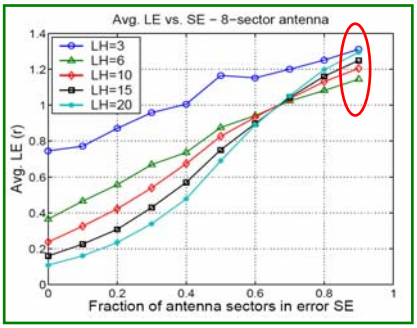
30

Performance Evaluation

- Localization Error vs. Sector Error
  - Sector error (SE): $SE = \frac{\#\text{ of sectors falsely estimated}}{LH}$
  - SeRLoc is resilient to sector error.
    - Why?

    → Due to the majority vote in determination of overlapping region
  - However, beyond a threshold (SE>0.7), Avg. LE increases with Avg. LH
    - Why?

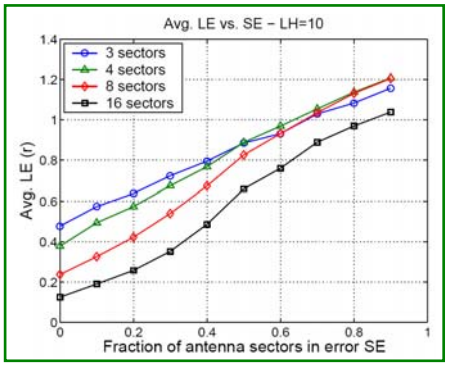
    → Falsely estimated sectors dominate in the location determination.



31

Performance Evaluation

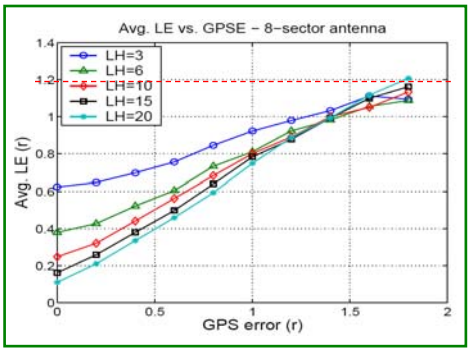
- Localization Error vs. Sector Error (continued)
  - Avg. LH = 10
  - The narrower the antenna sectors, the smaller the Avg. LE even in the presence of sector error.



32

Performance Evaluation

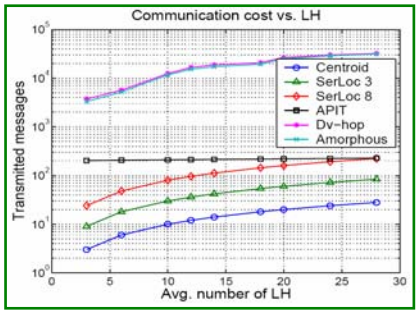
- Localization Error vs. GPS Error (GPSE)
  - GPSE = r means that every locator was randomly placed at a circle of radius r centered at the locator's actual position.
  - Even for the large GPSE, Avg. LE ≤ 1.2r.



33

Performance Evaluation

- Communication Cost vs. Locators Heard
  - DV-hop and Amorphous have significant higher comm. Cost.
    - Why? Due to the flood-based approach for the beacon propagation
  - APIT requires |L| + |N| beacons to localize the sensors.
  - SeRLoc requires |ML| number of beacons
  - Under assumption |N| >> |L|, SeRLoc has a smaller comm. than APIT.



34

Conclusion

- Proposed a range-independent, decentralized, localization scheme call SeRLoc.
- Also showed how the security mechanisms of SeRLoc combined with its inherent geometric properties.
- Provided accurate location estimation even in the presence of severe security threats in WSN, such as the wormhole and sybil attack.
- Showed through simulations that SeRLoc localizes sensors with higher accuracy than state-of-the-art range-independent localization schemes, while having lower communication cost.
- Moreover, showed that SeRLoc is resilient to sources of error such as error in the location of the reference points as well as error in the sector determination.

35