Robust Detection of Primary User Emulation Attacks in IEEE 802.22 Networks

Olga León
Department of Telematics
Universitat Politècnica de Catalunya (UPC)
Barcelona, Spain
olga@entel.upc.edu

Juan Hernández-Serrano
Department of Telematics
Universitat Politècnica de Catalunya (UPC)
Barcelona, Spain
jserrano@entel.upc.edu

Miguel Soriano
Universitat Politècnica de Catalunya (UPC)
& Centre Tecnològic de Telecomunicacions de Catalunya (CTTC)
Barcelona, Spain
soriano@entel.upc.edu

ABSTRACT

Cognitive Radio (CR) technology constitutes a new paradigm where wireless devices can access the spectrum left unused by licensed or primary users in an opportunistic way. This feature opens the door to a main new threat: the Primary User Emulation (PUE) attack, in which a malicious user transmits a fake primary signal preventing a Cognitive Radio Network (CRN) from using the available spectrum. Cooperative location of a primary source can be a valuable tool for distinguishing between a legitimate transmission and a PUE attack whenever the position of primary users is known, as it is the case of TV towers in the IEEE 802.22 standard. However, the location process can be undermined due to false data provided by malicious or faulty nodes. In this paper, we analyze the effect of forged reports on the location process of a given emitter and provide a set of countermeasures in order to make it robust to undesired behaviors.

Categories and Subject Descriptors

K.6 [**MANAGEMENT OF COMPUTING AND INFORMATION SYSTEMS**]: Security and Protection; D.2.8 [**COMPUTER-COMMUNICATION NETWORKS**]: General—*Security and protection*

General Terms

Security,Design

Keywords

CRN, PUE, security, localization

1. INTRODUCTION

Cognitive Radio Networks (CRNs) [2] are regarded to be a possible solution to the current underutilization of the spectrum by allowing Cognitive Radios (CRs) to act as secondary users of the spectrum left unused by licensed services. Thus, spectrum sensing is a crucial task in order to detect vacant bands or white spaces and avoid interfering primary transmissions. If a primary signal is detected in the operation channel, the CRN must switch to another band (a process known as spectrum handoff). On the other hand, if another secondary user is already operating in such band, self-coexistence mechanisms are needed to share the spectrum fairly.

With the specific characteristics of CRNs, new threats have arisen [5, 10]. In particular, the community research has paid special attention to the Primary User Emulation (PUE) attack and False Feedback attacks, since they can severely undermine the primary detection process. In the PUE attack, first coined in [3], an attacker pretends to be a primary user or incumbent by transmitting a signal with similar characteristics to a primary signal or replying a real one, thus preventing secondary users from using a vacant band. Consequently, there is a need for providing effective methods in order to distinguish between legitimate primary transmissions and fake ones (PUE attacks).

Research on this topic has been generally based on the recently approved standard IEEE 802.22 Wireless Regional Area Networks (WRANs) [1], that defines a centralized network composed by a Base Station (BS) and a set of CRs. In such kind of networks, two different types of primary users are defined: TV emitters and wireless microphones. Most of the proposals dealing with PUE attacks [9] in 802.22 networks rely on energy-based sensing mechanisms, and their performance is considerably reduced mainly due to shadowing. Localization techniques can be of paramount help in order to discriminate between real primary transmissions and PUE attacks [4]. According to the estimated position of the primary source, the CRN could decide whether the primary source is a TV primary transmitter, whose position is known, and, in any case, precisely locate the source.

In this paper, we analyze the effect of forged reports on the location-based PUE detection process of a CRN and provide a set of countermeasures in order to make this process robust to false feedback.

The structure of this paper is as follows. Sect. 2 provides an overview of the main techniques used for location of RF transmissions and sketches the current localization methods suitable for CRNs. In Sec. 3, the current threats to those

CogART '11 October 26-29, Barcelona, Spain

methods are identified and then analyzed. Next, in Sec. 4, some countermeasures are both presented and evaluated. Finally, Sec. 5 provides the conclusions of this work.

2. LOCALIZATION IN CRNS

Among the different existing location techniques for wireless networks [15], Received Signal Strength (RSS) and Time Difference of Arrival (TDoA) seem to be the most suitable for detecting PUE attacks. Opposite to Global Positioning System (GPS) or Time of Arrival (ToA), they do not require the cooperation of the node to be located, which cannot be expected from either and attacker or a real primary source.

With RSS-based techniques, assuming that the transmission power and the path loss model are known, it is possible to estimate the distance from the source to the reference node. When transmission power is not known, differences between RSS measured at pairs of receivers can be considered [12] removing in this way the dependency on the actual transmit power. A set of at least three RSS measurements is then used to estimate the position of the emitter by applying trilateration. Although RSS measurements are relatively inexpensive and simple to implement in hardware [15], they are susceptible of high errors due to the dynamics of indoor/outdoor environments mainly due to multipath signals and shadowing. The effect of shadowing is usually modeled as log-normal and the standard deviation of received power σ_{dB} leads to RSS-based estimates with variance proportional to its range. For this reason, it may not be suited for networks with long-range links, that is the case of WRAN 802.22 networks with ranges of order of kilometers.

On the other hand, TDoA is based on the difference of the time of arrival of a single signal (transmitted by the node to be located) at two different reference nodes. Note that, as these measures do not depend on the transmitter's clock, TDoA can be applied for locating asynchronous transmitters, as it is the case of a PUE attacker. As disadvantages, it requires a tight synchronization between each pair of reference nodes and TDoA measures are also quite sensitive to multipath propagation. However, TDoA provides considerably higher accuracy than RSS.

In TDoA, if a signal was received at time t_1 by the first reference node and reached the second reference node at t_2, the difference of distances Δd between the transmitter and both receivers is given by $v_p \cdot (t_1 - t_2)$, where v_p is the propagation velocity of the signal. When multiple TDoA measures are available, multilateration can be applied in order to estimate the node position. In a 2-dimensional space, each TDoA measurement defines a hyperbola on a surface and it is needed at least two TDoA measurements (three nodes or more) to locate an emitter, where the position is given by the intersection of both hyperbolas. On the other hand, in a 3-dimensional space, each measurement defines a hyperboloid and therefore three measures are required to locate the emitter. Nevertheless, in practice, measurements are subjected to errors and thus the different TDoA equations rarely intersect in a given solution. In this case, the location problem can be posed as an optimization problem and solved using, for example, a least squares (LS) method or an extended Kalman-Bucy filter. Since Kalman-Bucy provides no significant improvement in accuracy when the emitter to be located is usually stationary [7], LS methods over a linearized set of TDoA error equations (by means, for example, of Taylor-Series Estimations) are often preferred for stationary networks such as CRNs.

LS estimation methods [8] are iterative schemes that start with a rough initial guess (x_v, y_v, z_v) and improve the guess at each step $(x_v + \delta_x, y_v + \delta_y, z_v + \delta_z)$ by determining the local linear least-sum squared-error correction $(\delta_x, \delta_y, \delta_z)$. The target is to iterate the method until the components of the correction are below a given threshold, that is to say, that the estimation converges.

First, we have to obtain a linear estimation of the measurement errors. According to this, given a set of n TDoA measurements τ_i taken by the pairs made up of the BS and each one of the CRs, the measurement errors assuming a prediction (x_v, y_v, z_v) can be expressed as in (2), with $f_i(x, y, z)$ as in (1) the real TDoA measurement for the pair BS and anchor node i for position (x, y, z).

$$f_i(x,y,z) = \sqrt{(x-x_i)^2+(y-y_i)^2+(z-z_i)^2} - \sqrt{x^2+y^2+z^2} \quad (1)$$

$$\mathbf{e} = \begin{pmatrix} v_p\tau_1 - f_1(x_v,y_v,z_v) \\ v_p\tau_2 - f_2(x_v,y_v,z_v) \\ \vdots \\ v_p\tau_n - f_n(x_v,y_v,z_v) \end{pmatrix} \quad (2)$$

Then, from the 1st-degree Taylor polynomial of $\mathbf{e}$, the matrix representation of the linearized forms of the distance error can be expressed as in (3), with $\mathbf{A}$ an n-by-3 matrix with the Taylor coefficients and $\boldsymbol{\delta}$ a 3-by-1 column vector with the corrections $(\delta_x, \delta_y, \delta_z)$.

$$\hat{\mathbf{e}} = \mathbf{A}\boldsymbol{\delta} + \mathbf{e} \quad (3)$$

Assuming that $\hat{\mathbf{e}}$ is full rank, the value of $\boldsymbol{\delta}$ that minimizes the sum of quadratic errors $\hat{\mathbf{e}}^{\mathrm{T}}\hat{\mathbf{e}}$ can be computed as in (4).

$$\boldsymbol{\delta} = -(\mathbf{A}^{\mathrm{T}}\mathbf{A})^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{e} \quad (4)$$

However, in the real world, measurements performed by different nodes are subjected to different errors and then their measures may contribute to the LS estimation with different weights. Moreover, measurement errors are often correlated. Consequently, localization methods, instead of the previous approach, often minimize $\hat{\mathbf{e}}^{\mathrm{T}}\mathbf{W}\hat{\mathbf{e}}$, with $\mathbf{W}$ an n-by-n matrix with the assigned weights to every measure. In such case, the most common approach [7] is to define $\mathbf{W} = \mathbf{R}^{-1}$ with $\mathbf{R}$ the matrix of covariances between measures. Therefore, the optimal $\boldsymbol{\delta}$ can be derived as in (5).

$$\boldsymbol{\delta} = -[\mathbf{A}^{\mathrm{T}}\mathbf{W}\mathbf{A}]^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{W}\mathbf{e} \quad (5)$$

3. THREATS AND EFFECTS

Besides the usual threats to the localization method, that are more related to the variance of the acquired measures (SNR, multipath, fading, etc.), there are mainly two that could undermine the accuracy when localizing a primary source (real or fake): selfish nodes and liar nodes providing false feedback.

Selfish behavior reduces the amount of available measurements for the localization method and can become a big risk as the percent of selfish nodes increases. Nevertheless, selfish behavior in CRNs can be prevented with cooperation

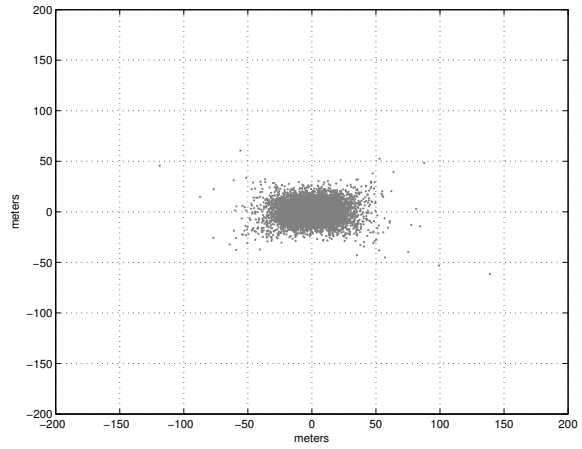

Figure 1: 2D prediction error of 10,000 iterations for locating a transmitter placed at (30000meters,0) in an 802.22 CRN with 30 collaborating CRs with an average reception SNR of -10dB

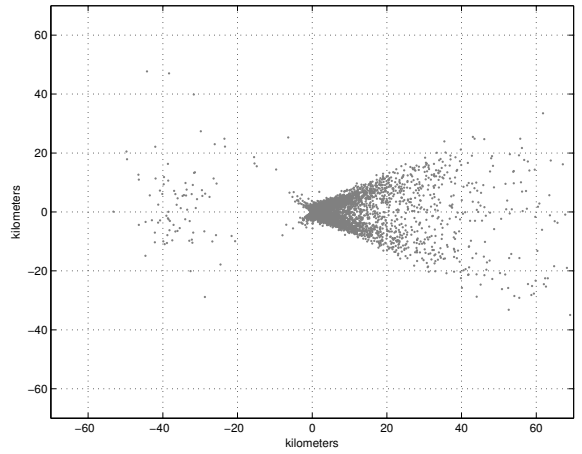

Figure 2: 2D prediction error in the same conditions as in Fig. 1 but with 2 CRs (liars) providing false feedback

enforcement mechanisms [13] without altering the localization method. For example, in 802.22 networks the BS can reduce or even cancel the bandwidth assigned to a given node if it does not collaborate.

On the other hand, the effect of liar nodes providing bad measures or false feedback is much challenging. Detecting and avoiding them or mitigating their effect on the localization method, as we will later explain in Sec. 4, requires changes to the cooperative localization method. In the following we will analyze the effect of false data on cooperative localization methods based on TDoA measurements, which are, as previously explained, the preferred localization methods when no collaboration can be expected from the source to be located.

Figures 1 and 2 are obtained considering an 802.22 network deployed over an square area of 60x60 km^2 composed by a BS located at the origin (0,0) and a set of 30 CRs nodes uniformly distributed within the area of the CRN. The attacker is placed at the CRN boundary at (30km,0). Whenever there is evidence of the existence of a primary transmission (the attacker), the BS requests the CRs to obtain TDoA measurements in order to locate the emitter. The obtained position will greatly help to effectively distinguish between a legitimate transmission and a PUE attack. The average SNR at the BS position is -10dB. Since there is no specific path loss model for the IEEE 802.22 standard being developed and the Okumura-Hata model has been widely used for UHF band measurements in digital TV reception [6], we have adopted it as the path loss model. For the location process, the Least Squares (LS) method described in Sec. 2 is applied in order to minimize the error performed in the estimation.

Figures 1 and 2 show the prediction error with 10,000 iterations of the localization method: the former when there are no liars; the latter when two liars report false feedback. We have adopted the worst case in which every liar reports a "credible" TDoA measure to its BS, that is to say, that provided measures are bounded to $\frac{d}{v_p}$, with d the distance between the liar and the BS, and v_p the propagation speed; measures above this bound would be discarded by any BS because they are not possible at all. Figures clearly show that the presence of just two liars out of 30 CRs results in a very inaccurate prediction (rising from order of meters to order of kilometers) that could often lead to false positives. This a devastating damage that, although it diminishes with the number of cooperating stations, brings to the arena the need for robust PUE localization methods.

4. ROBUST PUE DETECTION

As explained in Sec. 3, false reports provided by compromised nodes can severely undermine the location method thus leading to false positives or negatives regarding the detection of primary users. Consequently, there is a need for identifying false measurements in order to discard them for the location process. This task could be accomplished by comparing measurements from different nodes and looking for large deviations. However, measurements can considerably vary depending on the position of the CR within the CRN. Therefore, the most intuitive would be to group nodes into clusters and compare measurements among nodes belonging to the same cluster. Usually, outlier measurements may be (badly) detected by means of LS fitting, but we recommend, as some other researchers [11], to use Least Median Square (LMS) fitting instead. LMS aims to minimize the median of the residue squares as in (6) increasing its robustness to deviated measurements.

$$(x_v, y_v, z_v) = arg\,min\ [median_i\,(v_p\tau_i - f_i(x_v, y_v, z_v))] \quad (6)$$

However, the process of minimizing the median of the residue squares is prohibitive [17] and then the final position estimation should be obtained with a mixed solution:

1. Divide the set of n CRs into c several clusters of equal size $s = \lceil \frac{n}{c} \rceil$.

2. Apply the location process described in Sec. 3 separately in every cluster obtaining an estimation of the position of the emitter for each cluster (x_{v1}, y_{v1}, z_{v1})... (x_{vj}, y_{vj}, z_{vj})...(x_{vc}, y_{vc}, z_{vc}).

3. Compute the median of residue squares for each cluster

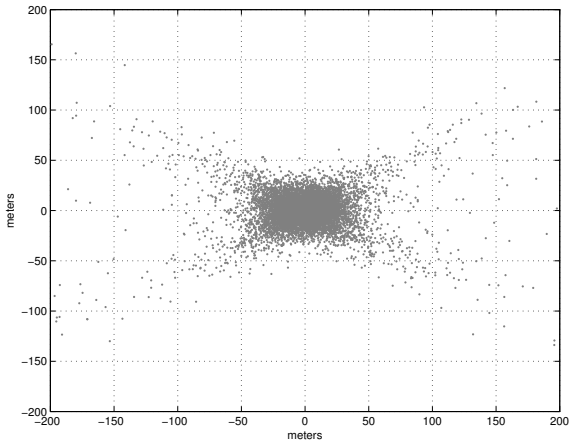

Figure 3: 2D prediction error in the same conditions as in Fig. 2 but with the CRN split into 3 equally-sized clusters

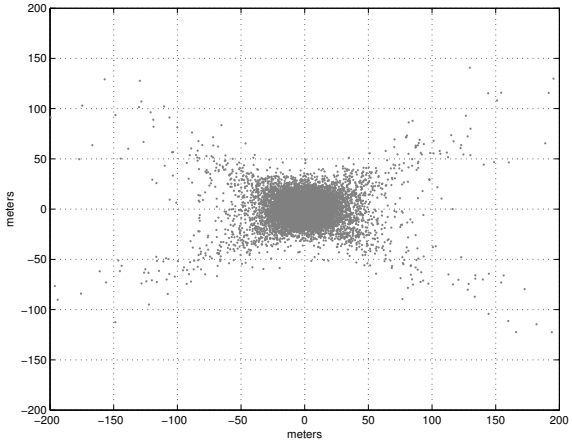

Figure 4: 2D prediction error in the same conditions as in Fig. 3 but with no liars providing false feedback

j as

$$r^2_{cluster_j} = median(r_1^2 ... r_i^2 ... r_s^2)$$

where $r_i = v_p\tau_i - f_i(x_{vj}, y_{vj}, z_{vj})$ is the residue for node i of cluster j and $f_i(x_{vj}, y_{vi}, z_{vi})$ as in (1) is a "error-free" TDoA measure for the position estimation obtained by means of LS method for cluster j.

4. Select as tentative estimation (x_v, y_v, z_v) the one given by the cluster with the lowest median of residues squares.
5. Compute the residue squares for all the n nodes considering the tentative estimation (x_v, y_v, z_v).
6. Perform a new position estimation by applying a LS method assigning a different weight to each node's measurement according to its residue square.

Assuming that at least in one of the clusters there are no false measurements, the LS estimation provided by the cluster will be more reliable and will exhibit a lower median of residues squares. Thus, when computing the residues in step 4 considering this estimation, false measurements will be clearly identified due to its higher value with respect to the rest. Since the accuracy of the location method improves as the number of (reliable) measurements increases, a final LS estimation is performed in the last step excluding outlier measurements or reducing their effect on the estimation by assigning a different weight to each measurement according to its residue. This is an implementation of Weighted Least Squares (WLS) method, in wich the estimated position $\boldsymbol{\delta} = (x_v, y_v, z_v)$ is obtained as in (5), defining W as a diagonal matrix with the weights assigned to each measurement as its diagonal elements w_{ii}.

Note that the election of the number of clusters plays an important role on the algorithm's performance. First, it should guarantee that there exists at least one cluster without liars; and second, when the number of stations n is small, the number of clusters should be also as small as possible to guarantee a minimum level of accuracy in the estimation performed by each cluster.

Fig. 3 depicts the error obtained in the position estimation with the same conditions as in Fig. 2 (two liars) when the set of CRs is divided into 3 clusters and the LMS method described above is applied. As it can be seen, the proposed mechanism can effectively identify deviated measurements and perform the location process almost with the same accuracy as in the case when there are no liars and a simple LS method is used.

Fig. 4 shows the effect of applying the LMS method when there are no liars. Once again, the accuracy of the prediction does not diminish with respect to the LS method although the error samples exhibit a slightly higher dispersion due to the weights mechanism used to ignore deviated measurements. Besides, it introduces some overhead with regard to the amount of computations performed in exchange for making the system robust to liars.

Finally, as compromised nodes are likely to report false data repeatedly, a trust mechanism should be integrated into the system so as to keep track of node's behavior over time. Trust and reputation models have been extensively studied specially in the context of ad hoc networks [14] and recently, the idea of applying them to CRNs to enhance collaborative spectrum sensing has recently attracted research interest [16]. In a similar way, they can be applied to the location process of an emitter in CRNs by weighting the measurements provided by each node according to the trust or reputation assigned by the system and computed considering not only the reliability of the current measurement but of those provided in the past.

5. CONCLUSIONS

CRNs appear as a promising solution to the scarcity of radio spectrum since they can "intelligently" select the best spectrum opportunities. However, their particular characteristics pose new security challenges. Among them, the main one is to provide mechanisms for distinguishing legitimate primary transmissions from PUE attacks without altering the primary network behavior. Research on this topic has been generally based on the IEEE 802.22 standard, in which two different types of primary users are defined: TV emitters and wireless microphones. Within these networks, localization of the transmission's source could help to provide valuable information to identify PUE attacks. In the case of a TV emitter, according to this estimation and the

real position of TV primary transmitters, which is assumed to be known to the CRN, the BS can take a decision about the legitimacy of the transmission. On the other hand, in the case of wireless microphones, the provided method precisely locates the source of emission.

Nevertheless, life is not a bowl of cherries, and the localization methods themselves are also subjected to several threats. In this paper, we have identified these threats and we have outlined the devastating impact of the most challenging one: malicious or compromised nodes providing false feedbacks. Moreover, we have outlined some solutions against these liars based on clustering and we have evaluated their strength. The presented results show that this is a promising research branch that is getting us to implement efficient and robust localization solutions. However, further research is still needed on the integration of reputation schemes with clustering; solutions not only providing accurate predictions of a primary source's location but also identifying malicious nodes.

6. ACKNOWLEDGEMENTS

This work has been partially supported by the Spanish *Comisión Interministerial de Ciencia y Tecnología* (CICYT) with the project P2PSEC (TEC2008-06663-C03-01), the Spanish *Ministerio de Ciencia e Innovación* with the CONSOLIDER project ARES (CSD2007-00004) and the *Generalitat de Catalunya* with the grant 2009 SGR-1362 to consolidated research groups awarded to the Information Security Group of the *Universitat Politècnica de Catalunya* (UPC).

7. REFERENCES

[1] IEEE Standard 802.22-2011 - Part 22: Cognitive Wireless RAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications: Policies and Procedures for Operation in the TV Bands, July 2011.

[2] I. F. Akyildiz, W.-Y. Lee, M. C. Vuran, and S. Mohanty. Next generation/dynamic spectrum access/cognitive radio wireless networks: a survey. *Comput. Netw.*, 50(13):2127–2159, 2006.

[3] R. Chen and J.-M. Park. Ensuring trustworthy spectrum sensing in cognitive radio networks. In *1st IEEE Workshop on Networking Technologies for Software Defined Radio Networks (SDR)*, pages 110–119, Sept. 2006.

[4] R. Chen, J.-M. Park, and J. Reed. Defense against primary user emulation attacks in cognitive radio networks. *Selected Areas in Communications, IEEE Journal on*, 26(1):25–37, Jan. 2008.

[5] T. Clancy and N. Goergen. Security in cognitive radio networks: Threats and mitigation. In *3rd International Conference on Cognitive Radio Oriented Wireless Networks and Communications (CrownCom)*, pages 1–8, May 2008.

[6] A. Eksim, S. Kulac, and M. Sazli. Effective cooperative spectrum sensing in IEEE 802.22 standard with time diversity. In *International Conference on Advances in Computational Tools for Engineering Applications, ACTEA'09.*, pages 528–531, July 2009.

[7] W. Foy. Position-location solutions by taylor-series estimation. *Aerospace and Electronic Systems, IEEE Transactions on*, AES-12(2):187 –194, Mar. 1976.

[8] T. C. Hsia. *System identification: Least-squares methods*. Lexington Books, 1977.

[9] Z. Jin, S. Anand, and K. P. Subbalakshmi. Mitigating primary user emulation attacks in dynamic spectrum access networks using hypothesis testing. *SIGMOBILE Mob. Comput. Commun. Rev.*, 13:74–85, Sept. 2009.

[10] O. León, J. Hernández-Serrano, and M. Soriano. Securing cognitive radio networks. *International Journal of Communication Systems*, 23(5):633–652, Feb. 2010.

[11] Z. Li, W. Trappe, Y. Zhang, and B. Nath. Robust statistical methods for securing wireless localization in sensor networks. In *Information Processing in Sensor Networks, 2005. IPSN 2005. Fourth International Symposium on*, pages 91–98. IEEE, 2005.

[12] C. Liu, T. Scott, K. Wu, and D. Hoffman. Range-free sensor localisation with ring overlapping based on comparison of received signal strength indicator. *Int. J. Sen. Netw.*, 2:399–413, July 2007.

[13] G. Marias, P. Georgiadis, D. Flitzanis, and K. Mandalas. Cooperation enforcement schemes for manets: a survey. *Wireless Communications and Mobile Computing*, 6(3):319–332, 2006.

[14] M. Mejia, N. Pena, J. L. Munoz, and O. Esparza. A review of trust modeling in ad hoc networks. *Internet Research*, 19(1):88–104, 2009.

[15] N. Patwari, J. Ash, S. Kyperountas, I. Hero, A.O., R. Moses, and N. Correal. Locating the nodes: cooperative localization in wireless sensor networks. *Signal Processing Magazine, IEEE*, 22(4):54–69, July 2005.

[16] T. Qin, H. Yu, C. Leung, Z. Shen, and C. Miao. Towards a trust aware cognitive radio architecture. *ACM SIGMOBILE Mobile Computing and Communications Review*, 13(2):86–95, 2009.

[17] P. Rousseeuw, A. Leroy, and J. Wiley. *Robust regression and outlier detection*, volume 3. Wiley Online Library, 1987.