Lecture Notes in Networks and Systems 34

Basant Tiwari · Vivek Tiwari
Kinkar Chandra Das
Durgesh Kumar Mishra
Jagdish C. Bansal *Editors*

Proceedings of International Conference on Recent Advancement on Computer and Communication

ICRAC 2017

Springer

Lecture Notes in Networks and Systems

Volume 34

Series editor

Janusz Kacprzyk, Polish Academy of Sciences, Systems Research Institute, Warsaw, Poland
e-mail: kacprzyk@ibspan.waw.pl

The series "Lecture Notes in Networks and Systems" publishes the latest developments in Networks and Systems—quickly, informally and with high quality. Original research reported in proceedings and post-proceedings represents the core of LNNS.

Volumes published in LNNS embrace all aspects and subfields of, as well as new challenges in, Networks and Systems.

The series contains proceedings and edited volumes in systems and networks, spanning the areas of Cyber-Physical Systems, Autonomous Systems, Sensor Networks, Control Systems, Energy Systems, Automotive Systems, Biological Systems, Vehicular Networking and Connected Vehicles, Aerospace Systems, Automation, Manufacturing, Smart Grids, Nonlinear Systems, Power Systems, Robotics, Social Systems, Economic Systems and other. Of particular value to both the contributors and the readership are the short publication timeframe and the world-wide distribution and exposure which enable both a wide and rapid dissemination of research output.

The series covers the theory, applications, and perspectives on the state of the art and future developments relevant to systems and networks, decision making, control, complex processes and related areas, as embedded in the fields of interdisciplinary and applied sciences, engineering, computer science, physics, economics, social, and life sciences, as well as the paradigms and methodologies behind them.

Advisory Board

More information about this series at http://www.springer.com/series/15179

Basant Tiwari · Vivek Tiwari
Kinkar Chandra Das · Durgesh Kumar Mishra
Jagdish C. Bansal
Editors

Proceedings of International Conference on Recent Advancement on Computer and Communication

ICRAC 2017

Springer

Editors
Basant Tiwari
Computer Science and Engineering
Technocrats Institute of Technology
Bhopal, Madhya Pradesh
India

Vivek Tiwari
Computer Science and Engineering
DSPM IIIT
Naya Raipur, Chhattisgarh
India

Kinkar Chandra Das
Department of Mathematics
Sungkyunkwan University
Suwon
Korea (Republic of)

Durgesh Kumar Mishra
Microsoft Innovation Academy, Computer Science and Engineering
Sri Aurobindo Institute of Technology
Indore, Madhya Pradesh
India

Jagdish C. Bansal
Department of Mathematics
South Asian University
New Delhi, Delhi
India

ISSN 2367-3370 ISSN 2367-3389 (electronic)
Lecture Notes in Networks and Systems
ISBN 978-981-10-8197-2 ISBN 978-981-10-8198-9 (eBook)
https://doi.org/10.1007/978-981-10-8198-9

Library of Congress Control Number: 2018930372

Printed on acid-free paper

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd. part of Springer Nature
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

Preface

On May 25–26, 2017, the iMPLab Research & Innovation at Bhopal organized the first International Conference on Recent Advancement on Computer and Communication (ICRAC 2017).

In an era of sensor, high-speed wireless communication, energy-efficient device, huge data, information security, ubiquitous computing, and Internet of Things (IoT) are undergoing considerable change. The rapid influx of digital technologies is redefining communication processes, altering media structures and operations, and transforming the availability and accessibility of information. While these developments enable new social, cultural, political, and economic activities, they can also produce adverse ramifications for consumers and communities in terms of all aspects of security.

Overall, the conference hosted nearly 82 presentations by academicians from India and abroad. The speakers covered a broad range of issues and challenges in the Internet and telecommunications fields including the influence of IoT application to make things smarter; digital infrastructure developments and opportunities for further connectivity, particularly in relation to mobile and public Wi-Fi prospects; new methods for information security; advancing communication practices to facilitate social inclusion; energy-efficient sensors and other devices; and the other challenges associated with changing technologies and communication.

The papers published in the volume represent an overview of the issues explored at the conference and accentuate the diversity of both research topics and methodologies used in the communication. Moreover, they reflect salient, nuanced, and interrelated emerging issues surrounding information flow, digital participation, IoT. These papers have undergone double-blind peer review, with Indian and international academics assisting in this process. Many thanks to all of the authors for the time and effort spent developing papers for the proceedings. Hosting the International Conference on Recent Advancement on Computer and Communication first year was no small feat and certainly would not have been possible without the broad support it received. In particular, sincere thanks to all our friends, colleagues, mentors, and research scholars who helped out each day. In addition to people from industry and academia helping to promote it, assisting with the peer review process, chairing sessions, presenting research, contributing to the proceedings, and engaging with speakers in the sessions—we would not have been able to hold the event without your support and greatly appreciate the efforts and willingness of all who were involved.

Bhopal, India — Basant Tiwari
Naya Raipur, India — Vivek Tiwari
Suwon, Korea — Kinkar Chandra Das
Indore, India — Durgesh Kumar Mishra
New Delhi, India — Jagdish C. Bansal

Contents

About the Editors

Dr. Basant Tiwari is currently serving as Professor and Head in Computer Science and Engineering Department at Technocrats Institute of Technology, Bhopal. He has rich experience in teaching undergraduate and postgraduate classes. He has many international and national publications to his credit in conferences and journals. He has also attended many national and international conferences, workshops, seminars, and symposiums all over India and abroad. His current area of research is "Pervasive Healthcare and Remote Medical Care." He is a Senior Member of IEEE, ACM, CSI, and IACSIT. Currently, he has been honored as Bhopal Representative in IEEE M.P. Sub-section, Treasurer in ACM Udaipur Chapter, and Secretary in CSI Bhopal Chapter. He has organized various national and international conferences, delivered invited talks, and also chaired the technical sessions. He is a reviewer of various reputed international journals and books. He did his Ph.D. from Devi Ahilya University, Indore.

Dr. Vivek Tiwari is an Assistant Professor of Computer Science and Engineering at DSPM International Institute of Information Technology, Naya Raipur, India. He is alumni of Maulana Azad National Institute of Technology (MANIT), Bhopal. Previously, he worked with the universities of national repute and Caresoft Incorporation (based at Middlesex, NJ, USA) as Software Engineer. He is the recipient of Young Scientist Fellowship (MPYSC_2014_814) for the year 2014–2016 by the Madhya Pradesh Council of Science & Technology (MPCST), Government of Madhya Pradesh. He has published more than 30 research papers, book, and chapters in the areas of data mining, data warehousing, pattern warehousing, distributed computing, and cloud computing in leading international journals (Springer, Inderscience, Elsevier, ACM, and IGI Global) and conferences. He is editor in chief of book "Handbook of Research on Pattern and Data Analysis in Healthcare Settings" under the series of Advances in Data Mining & Database Management (ADMDM), published by IGI Global, USA. A research project

"Aakash for Education" of cost 5 lakh funded by MHRD, India, is in his credit. He is associated with National Mission on Education through ICT (NMEICT) supported by IIT Bombay and MHRD since 2012.

Dr. Kinkar Chandra Das obtained his M.Tech. degree in Computer Science and Data Processing and Ph.D. degree in Spectral Graph Theory from Indian Institute of Technology Kharagpur, India, in 2004. He won the French Scholarship from Ministry of France, helping him to be a part of LRI, University Paris XI, France, for 1 year. After that, he joined the Department of Mathematics, Sungkyunkwan University, in 2006, and presently, he holds the position of Associate Professor in the same university. His main areas of research interest are spectral graph theory, molecular graph theory, degree sequence of graphs, and graph coloring. He has published almost 200 research papers in these areas in reputed international journals, guided two Ph.D. students, and supervised five postdoctoral researchers, and four Ph.D. students are working at the moment. He has written a book on "The Harary Index of a Graph" with two coauthors, published by Springer, Heidelberg, Germany, in 2014. He has recently edited three monographs on Mathematical Chemistry: (i) Bounds in Chemical Graph Theory—Basics, (ii) Bounds in Chemical Graph Theory—Mainstreams, and (iii) Bounds in Chemical Graph Theory—Advances. He has been on the editorial boards of various reputed journals.

Dr. Durgesh Kumar Mishra has received his M.Tech. degree in Computer Science from DAVV, Indore, in 1994 and Ph.D. degree in Computer Engineering in 2008. Presently, he has been working as a Professor (CSE) and Director, Microsoft Innovation Centre, Sri Aurobindo Institute of Technology, Indore, MP, India. He is also a visiting faculty at IIT Indore, MP, India. He has 24 years of teaching and 10 years of research experience. He has completed his Ph.D. under the guidance of late Dr. M. Chandwani on Secure Multi-Party Computation for Preserving Privacy. He has published more than 90 papers in refereed international/national journals and conferences including IEEE and ACM. He visited and delivered his invited talk in Taiwan, Bangladesh, Singapore, Nepal, USA, UK, and France. He has authored a book on "Database Management Systems." He had been Chief Editor of Journal of Technology and Engineering Sciences.

Dr. Jagdish C. Bansal is an Assistant Professor at South Asian University, New Delhi, India. Holding an excellent academic record, he is an outstanding researcher in the field of swarm intelligence at the national and international levels, having written several research papers in journals of national and international repute.

Impact of Various Networks Security Attacks on Wireless Sensor Localization Algorithms Based upon WSN Node's Residual Energy

Santosh Soni and Manish Shrivastava

Abstract Presently, wireless sensor network localization algorithms attracted various researchers toward research study and experiments. The location of nodes can be determined by various localization algorithms. These wireless sensor network localization algorithms are vulnerable and can be compromised for their security. There are various network security attacks like wormhole, impersonation, compromise, and duplicate attacks which degrade the performance of WSNs. In this research paper, under various network security attacks, we have tested wireless sensor network localization algorithms under certain performance matrices like mobility of WSN node's, size of packets, temperature, and node density to find out the compromised WSN node's residual energy. The simulation result shows how these localization algorithms are vulnerable to network security attacks.

Keywords Wireless sensor network · WSN localization algorithms
Network security attacks

1 Introduction

In the defense, civil, and critical types of fields, wireless sensor networks contain a lot of various features like autonomous system which can be installed in various locations and able to perform secure gathering of data, surveillance or monitoring the desired location. In this type of situations, the main aspect is security of wireless sensor network [1]. The localization plays key role in wireless sensor network and having

S. Soni (✉)
Department of IT, Institute of Technology,
Guru Ghasidas Vishwavidyalaya, Bilaspur, Chhattisgarh, India
e-mail: santoshsoni.77@gmail.com

M. Shrivastava
Department of CSE, Institute of Technology,
Guru Ghasidas Vishwavidyalaya, Bilaspur, Chhattisgarh, India
e-mail: manbsp@gmail.com

B. Tiwari et al. (eds.), *Proceedings of International Conference on Recent Advancement on Computer and Communication*, Lecture Notes in Networks and Systems 34, https://doi.org/10.1007/978-981-10-8198-9_1

vulnerabilities, this system may become target for any network security attacks. In this research study, we have shown that presently available localization algorithms are vulnerable to these network security attacks. The impact of network security attacks finally reduces the residual energy of wireless sensor network's node. Mainly, the gathered data and sensor motes should be positioned in space to find the exact location of an event. This position-based system is able to get execute using a localization-based approach. Localized systems are really critical issue of WSNs, because they aim to find events as well as used as the main station for routing, controlling density, tracking, and some other protocols [2]. Following are the various section of this research study:

Section 1 Introduction	Section 2 Previous Work
Section 3 WSN Localization Algorithms	Section 4 Network Security Attacks
Section 5 Simulation Setup and Results	Section 6 Conclusion and Future Work

2 Previous Work

Various research studies have been taken into account to consider the research outcome based upon certain performance matrices. Some of them are as follows:

[I] Azzedine Boukerche, Horacio A. B. F. Oliveira, Eduardo F. Nakamura, Antonio A. F. Loureiro [2] have studied and described localization system under various network security attacks with various security techniques.

[II] Ismail Güvenc., Member, Chia-Chin Chong [3] have shown various numerous localization algorithms along with certain accuracies.

[III] Nick Iliev, Igor Paprotny [4] have provided a complete review and implementations.

[IV] Svarika Goyal, Tarunpreet Bhatia, A. K. Verma [5] In this paper, different attacks are studied along with various defensive techniques.

[V] T. Meena, M. Nishanthy, Mr. E. Kamalanaban [6] proposed how to detect spoofing attacks.

[VI] Xu Huang, Muhammad Ahmed, Dharmendra Sharma [7] described various cryptographic methods defend against various network security attacks.

[VII] Asma Mesmoudi, Mohammed Feham, Nabila Labraoui [8] In general, localization techniques are classified with two types of algorithms: range-based and range-free.

3 WSN Localization Algorithms

Localization systems, which give critical location information of sensors and event occurrences, which can be the point of attacks can lead to compromise the working of a WSN [9]. There are various WSN localization algorithms like KALMAN, MCL, IMCL, MPL, and SMPL. Localization schemes usually focus on static sensor networks where the sensor nodes do not move once they are deployed, and sensor nodes. The issue of location indicates the method of searching location for sensor motes with specified coordinate techniques. For the localization of a sensor network with the GPS, certain specific motes shall be known of their place of installation mainly from already developed system or placing manually, which are known anchors (beacons). Other ones, which are specifically called unknown nodes generally used to get installed place through certain algorithms [10]. This contains of two steps: (a) measure of distance among motes and (b) typical measurement methods. Using the distance measurement methods, localization algorithms are fallen between range and free of range methods. The algorithms can be estimated based on both the coordinate estimation errors and the distance estimation errors [11, 12]. This is presumed that distance measurements between given pair of nodes that are exists within the radio range [12], which is

$$d(i,j) = d(i,j) + \varepsilon ij, \quad 1 \leq i, j \leq N, i \neq j \tag{1}$$

where N is node from network, ij is distance estimation error, and

$$d(i,j) = \sqrt{}(xi - xj)^{\wedge}2 + \sqrt{}(yi - yj)^{\wedge}2 \tag{2}$$

where (xi, yi) and (xj, yj) are the coordinates of nodes i and j, respectively. All the systems discussed in that research need a complete setup and are equipped with toward location sensing for office applications than for sensor nodes. That is why we cannot allow them in truly ad hoc WSNs [13]. Below are the two methods for sensor mote's installed place:

(a) Relative location: This is better shown by given formulas [3]:

$$\left.\begin{aligned} \sqrt{}(x - xa)^{\wedge}2 + (y - ya)^{\wedge}2) = da \\ \sqrt{}(x - xb)^{\wedge}2 + (y - yb)^{\wedge}2) = db \\ \sqrt{}(x - xc)^{\wedge}2 + (y - yc)^{\wedge}2) = dc \end{aligned}\right\} \tag{3}$$

(b) Triangulation: This calculates the exact position of unknown node which is depending on the angular distance among three various pairs of anchor nodes represented by below equations [4].

$$\left.\begin{aligned} \sqrt{}(xo1 - xa)^{\wedge}2 + (yo1 - ya)^{\wedge}2) = r1 \\ \sqrt{}(xo1 - xb)^{\wedge}2 + (yo1 - yb)^{\wedge}2) = r1 \\ \sqrt{}(Xa - xc)^{\wedge}2 + (ya - yc)^{\wedge}2) = 2r1^{\wedge}2 - 2r1^{\wedge}2 \cos \alpha \end{aligned}\right\} \tag{4}$$

4 Network Security Attacks

Security attacks which represent distance and position computations are very specific attacks in existing localized systems. Moreover, the various parts of a localization system, develops the equal type of vulnerabilities connected along towards various systems, since it is distributed multihop algorithm. For example, these below attacks include Sybil, replay, wormhole along with duplicate attack [2, 5–7, 14].

Sybil Attack: Here, a malicious node is a set of different node and continues sending wrong information. Such wrong information can be distance, positions estimation, various hops, or beacons.

Replay Attack: Here, from beacon node, a compromised node copies a received packet and starts resending the same packet after some time. This is the clone of the initial packet, The neighboring nodes wrongly deduct that the infected node is the node which has sent out the initial packet. Here, distance computation is obtained on the basis of compromised node where the final position in the packet highly depends on the initial node, and then, position computation gets affected. Mainly, signal strength with time-based distance calculations are affected, since the packets sent by the compromised node will have a totally discrete signal strength as well as discrete propagation time [2, 5–7, 14].

Wormhole Attack: Here, the information obtained by single infected node is sent to other side of the network and cloned by other infected node on the different side of the network. The given path between mentioned two attackers is a wormhole in the manner that a packet coming on one side is copied and forwarded on the different side of the network, representing as if it came from different neighboring node. Such kind of attack is shown in. Finally, these attacks can deeply destroy an insecure localization system through setting completely different and erroneous reference points in the position computations [2, 5–7, 14].

Compromise Attack: This is having three stages: physically compromising the sensors, deploying the compromised sensors, and compromised nodes after starting attacks [1, 15].

5 Simulation Setup and Result

WSN localization estimation is the process of getting the final position of sensor nodes; such position is based upon estimation but not right. Because GPS got high cost and calibrated outdoors only, here some localization algorithm looks to be perfect in calculating the exact position of sensor nodes [16]. The used WSN localization simulation tool is based upon two layers, a core simulator layer and a localization layer. Here, Table 1 presents various WSN localization simulation parameters information:

Simulation Results: Simulation results are specified in Table 1. The graph in Figs. 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12 and Tables 2, 3, 4, 5 show the performance of localization algorithms in form of WSN node's residual energy against various network security attacks.

Table 1 WSN localization simulation parameters

Simulator	WSN Localization Simulator
WSN localization Algorithms	Kalman, Monte Carlo, improve Monte Carlo, mobility prediction localization, secure MPL
Network security attacks	Wormhole, Sybil, compromise and replay
Mobility of WSN nodes	10, 20, and 30 m/s
Packet size	256, 512, and 1024 Bytes
Temperature	20°, 25°, and 30°
WSN node density	50, 100, and 150
Sensor's mobility model	Modified random waypoint
Anchor's mobility model	Modified random waypoint
Number of anchor nodes	150
Sensor model	Mica2
Propagation model	Two-ray ground
Simulation time	150 s

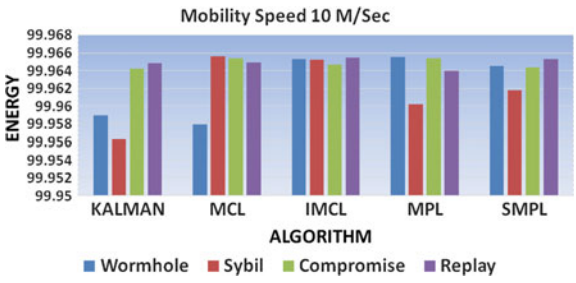

Fig. 1 Performance of WSN localization algorithms versus network security attacks under mobility speed 10 M/S

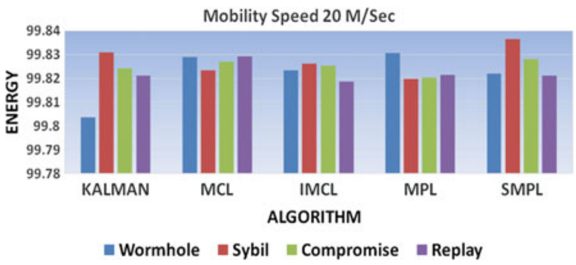

Fig. 2 Performance of WSN localization algorithms versus network security attacks under mobility speed 20 M/S

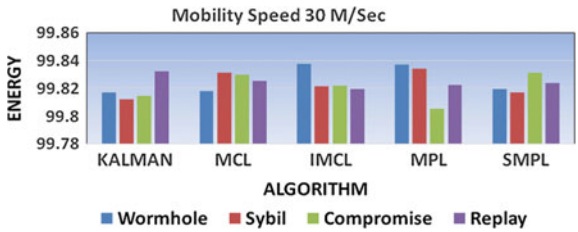

Fig. 3 Performance of WSN localization algorithms versus network security attacks under mobility speed 30 M/S

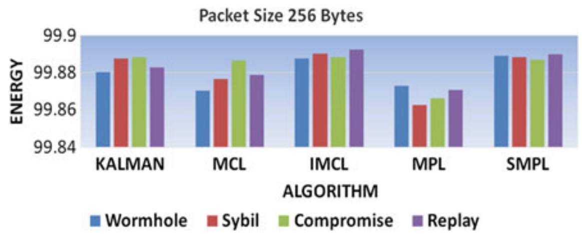

Fig. 4 Performance of WSN localization algorithms versus network security attacks under packet size 256 Bytes

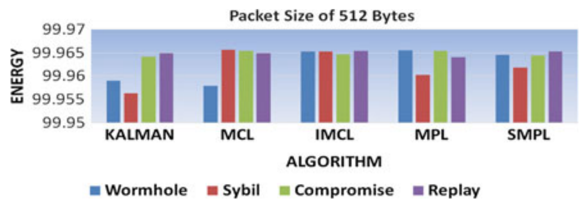

Fig. 5 Performance of WSN localization algorithms versus network security attacks under packet size 512 Bytes

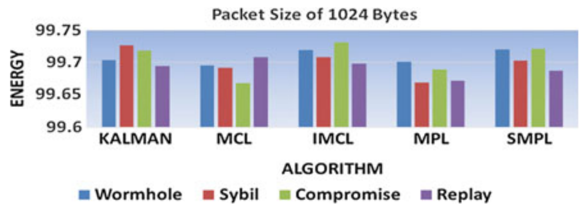

Fig. 6 Performance of WSN localization algorithms versus network security attacks under packet size 1024 Bytes

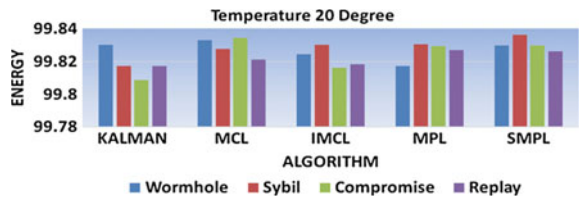

Fig. 7 Performance of WSN localization algorithms versus network security attacks under 20° temperature

Fig. 8 Performance of WSN localization algorithms versus network security attacks under 25° temperature

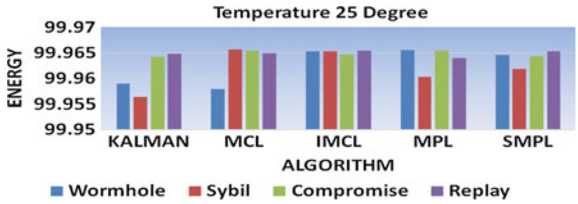

Fig. 9 Performance of WSN localization algorithms versus network security attacks under 30° temperature

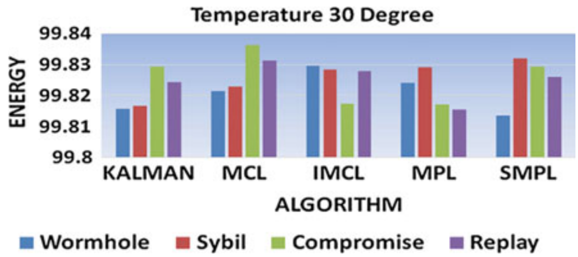

Fig. 10 Performance of WSN localization algorithms versus network security attacks with the node density of 50 nodes

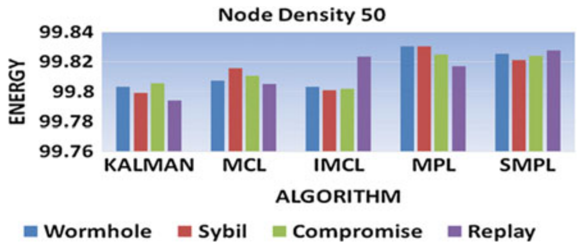

Fig. 11 Performance of WSN localization algorithms versus network security attacks with the node density of 100 nodes

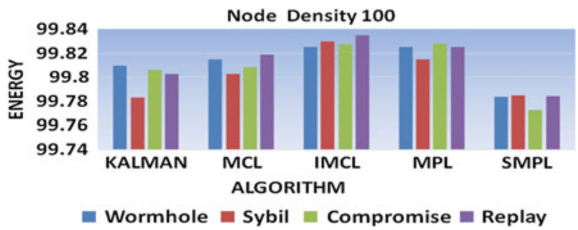

Fig. 12 Performance of WSN localization algorithms versus network security attacks with the node density of 150 nodes

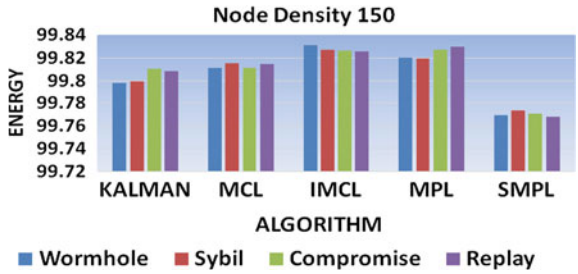

Table 2 Vulnerability of WSN localization algorithm against various network security attacks under the performance matrices of mobility of WSN nodes

	Wormhole	Sybil	Compromise	Replay
1. Kalman	Vulnerable	Vulnerable	Not vulnerable	Not vulnerable
2. MCL	Vulnerable	Not vulnerable	Not vulnerable	Not vulnerable
3. IMCL	Not vulnerable	Not vulnerable	Not vulnerable	Vulnerable
4. MPL	Not vulnerable	Vulnerable	Vulnerable	Not vulnerable
5. SMPL	Not vulnerable	Vulnerable	Not vulnerable	Vulnerable

Table 3 Vulnerability of WSN localization algorithm against various network security attacks under the performance matrices of packet size of WSN nodes

	Wormhole	Sybil	Compromise	Replay
1. Kalman	Vulnerable	Vulnerable	Not vulnerable	Vulnerable
2. MCL	Vulnerable	Not vulnerable	Vulnerable	Not vulnerable
3. IMCL	Vulnerable	Not vulnerable	Vulnerable	Vulnerable
4. MPL	Not vulnerable	Vulnerable	Not vulnerable	Not vulnerable
5. SMPL	Not vulnerable	Vulnerable	Vulnerable	Vulnerable

Table 4 Vulnerability of WSN localization algorithm against various network security attacks under the performance matrices of temperature

	Wormhole	Sybil	Compromise	Replay
1. Kalman	Vulnerable	Vulnerable	Vulnerable	Not Vulnerable
2. MCL	Vulnerable	Not vulnerable	Not vulnerable	Vulnerable
3. IMCL	Not vulnerable	Not vulnerable	Vulnerable	Not vulnerable
4. MPL	Vulnerable	Vulnerable	Not vulnerable	Vulnerable
5. SMPL	Vulnerable	Vulnerable	Not vulnerable	Vulnerable

Table 5 Vulnerability of WSN localization algorithm against various networks security attacks under the performance matrices of WSN node density

	Wormhole	Sybil	Compromise	Replay
1. Kalman	Vulnerable	Vulnerable	Not vulnerable	Vulnerable
2. MCL	Not vulnerable	Vulnerable	Vulnerable	Vulnerable
3. IMCL	Vulnerable	Vulnerable	Not vulnerable	Vulnerable
4. MPL	Not vulnerable	Vulnerable	Not vulnerable	Vulnerable
5. SMPL	Not vulnerable	Vulnerable	Vulnerable	Vulnerable

6 Conclusion and Future Scope

Finally, this research study has shown the nature of vulnerability and impact of WSN localization algorithms which is summarized in Tables 2, 3, 4, and 5. Mainly, localization algorithms show basic help and support toward various location-based protocols along with different implementation areas. Accuracy of localization is deeply integrated to the QoS of WSNs [17]. In a nutshell, we can say that Kalman algorithm is vulnerable to wormhole and Sybil attacks. MCL, IMCL, and SMPL algorithms show good performance against Sybil attack. Now researchers can use these findings to develop/improve more secure WSN localization algorithms against various network security attacks. MPL algorithm shows good performance against compromise attack. In the future, we will study the security techniques/methods to better improve the further performance of localization algorithms against network security attacks.

References

1. Nema, A., Tiwari, B., Tiwari, V.: Improving accuracy for intrusion detection through layered approach using support vector machine with feature reduction. In: Proceedings of the ACM Symposium on Women in Research 2016, pp. 26–31, ACM (2016)
2. Boukerche, A., Oliveira, H.A.B.F., Nakamura, E.F., Loureiro, A.A.F.: Secure localization algorithms for wireless sensor networks. IEEE Commun. Mag. **08**, 96–101 (2008)
3. Güvenc, I., Chong, C.-C.: A survey on TOA based wireless localization and NLOS mitigation techniques. IEEE Commun. Surv. Tutor. **11**(3), 107–124 (2009)
4. Nick, I., Paprotny, I.: Review and comparison of spatial localization methods for low power wireless sensor networks. IEEE Sens. J. **15**(10), 5971–5987 (2015)
5. Goyal, S., Bhatia, T., Verma, A.K.: Wormhole and Sybil attack in WSN: a review. In: 2015 2nd International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, pp. 1463–1468, INDIACom (2015)
6. Meena, T., Nishanthy, M., Kamalanaban, E.: Cluster-based mechanism for multiple spoofing Attackers in WSN. In: International Conference on Information Communication and Embedded Systems (ICICES2014), Chennai, pp. 1–5 (2014)
7. Huang, X., Ahmed, M., Sharma, D.: Protecting from inside attacks in wireless sensor networks. In: IEEE Ninth International Conference on Dependable, Autonomic and Secure Computing, Sydney, NSW, pp. 186–191, IEEE, New York (2011)
8. Mesmoudi, Asma, Feham, Mohammed, Labraoui, Nabila: Wireless sensor networks localization algorithms: a comprehensive survey. Int. J. Comput. Netw. Commun. (IJCNC) **5**(6), 45–64 (2013)
9. Yang, W., Zhu, W.T.: Voting-on-grid clustering for secure localization in wireless sensor networks. IEEE ICC 2010 Proceedings, pp. 01–05, IEEE, New York (2010)
10. Hong, S.H., Kim, B.K., Eom, D.S.: Localization algorithm in wireless Sensor networks with network mobility. IEEE Trans. Consum. Electron. **55**(4), 1921–1928 (2009)
11. Dai, H., Chen, A.G., Gu, X.F., He, L.: Localization algorithm for large-scale and low-density Wireless sensor networks. Electron. Lett. **47**(15), 881–883 (2011)
12. Yu, K., Guo, Y.J.: Anchor-free localization algorithm and performance analysis in wireless sensor Networks. IET Commun. **3**(4), 549–560 (2009)

13. Zhang, Y., Wu, W., Chen, Y.: A range based localization algorithm for wireless sensor networks. J. Commun. Netw. **7**(4), 429 (2005)
14. Shaukat, H.R., Hashim, F., Sali, A., Fadlee Abdul Rasid, M.: Node replication attacks in mobile wireless sensor network: a survey. Int. J. Distrib. Sens. Netw. **2014(**Article ID 402541), 01–15 (2014)
15. Lin, Z.-T., Qu, Y.-G., Jing, L., Zhao, B.-H.: AP Web Workshops 2006 Compromised Nodes in Wireless Sensor Network, pp. 224–230, Springer, New York (2006)
16. https://www.codeproject.com/articles/606364/wireless-sensor-networkLocalization-simulator-v
17. Han, G., Jiang, J., Zhang, C., Duong, T.Q., Guizani, M., Karagiannidis, G.K.: A survey on mobile anchor node assisted localization in wireless sensor networks. IEEE Commun. Surv. Tutor. **18**(3), 2220–2243 (2016)
18. Chen, L.: Distributed range-free localization algorithm for Wireless Sensor networks. Electron. Lett. **50**(12), 894–896 (2014)
19. Lin, X.: CAT: building couples to early detect node compromise attack in wireless sensor networks. In: GLOBECOM 2009—2009 IEEE Global Telecommunications Conference, Honolulu, HI, pp. 1–6, IEEE, New York (2009)
20. Messai, M.-L.: Classification of attacks in wireless sensor networks. In: International Congress on Telecommunication and Application'14 University of A.MIRA Bejaia, Algeria, pp. 23–24 (2014)
21. Bharath Srinivas, V.: Spoofing attacks in wireless sensor networks. IJCSET **3**(6), 201–210 (2013)

Fitting a Neural Network Classification Model in MATLAB and *R* for Tweeter Data set

Syed Muzamil Basha and Dharmendra Singh Rajput

Abstract Nowadays, the interest among the research community in sentiment analysis (SA) has grown exponentially. Our paper aims to find the prediction error occurred when we perform SA on tweets. The data set considered for the demonstration has 1129 tweets, and output parameters having predictor identifiers. Artificial neural networks (ANNs) are designed with ten hidden layers and one output layer. Additionally, trained the designed system with the help of MATLAB software to find the prediction error and also, derived sentiments using ggplot2 package in *R*.

Keywords SA · ANN · ggplot2 · *R*

1 Introduction

This instruction file for *A* classifier *C* is a mapping, $C : R_n \rightarrow N_{\mathrm{Lc}}$, where R_n is multidimensional space and N_{Lc} is the set of vectors with different labels assigned for a specific class (C) problem and is termed as

$$N_{\mathrm{Lc}} = y \in R_c; \quad y_i \in \{0, 1\} \forall_i, \quad \sum_{i=1}^{c} y_i = 1.$$

By default, the nature of all the traditional classification algorithms is binary, which can be further enhanced to address multi-class problems. The goal of classification is to accurately predict the target class for each case in the data having 134 tweets as training data collected from http://www.sananalytics.com/lab/twitter-sentiment. This data set contains 5513 human-categorized reviews based on topic in the form of

S. M. Basha (✉) · D. S. Rajput
Vellore Institute of Technology University, Vellore 632014, India
e-mail: muza.basha@gmail.com

D. S. Rajput
e-mail: dharmendrasingh@vit.ac.in

B. Tiwari et al. (eds.), *Proceedings of International Conference on Recent Advancement on Computer and Communication*, Lecture Notes in Networks and Systems 34, https://doi.org/10.1007/978-981-10-8198-9_2

tweets. Each entry contains tweet id, tweet text, tweet creation date, topic used for sentiment, sentiment label: *{positive (p), Neutral (N), Negative (Ne), Others (O)}*. The entire classification of reviews was done by an American male who is fluent in English. Niek J. Sanders njs@sananalytics.com, http://linkedin.com/in/niekjsanders.

2 Related Work

The author [1] has proposed an algorithm, addressing multi-class problems by considering seven different data sets and made an extensive comparison of time and accuracy in training the model with classical Kozas model, whereas the author in [2] also addressed the multi-class problem in the application area called ConveNets with all possible scenarios. In [3], the author proposed a new version of convolution neural network multi-scale spatial partition network, which can handle both text and image data. In [4], the author has proposed a classifier which can achieve fine-granularity aesthetic quality prediction with an accuracy of 87.1%. In [5], the author made an attempt to improve the classification performance of fuzzy minimum and maximum algorithm using hyper-box expansion rule. In [6], author proposed deep neural network system (DNN) that uses hidden Markov model toolkit to generate transformed MFCC features from BNF extractor. In [7], the author addresses the problem of feature selection from noisy data by combining two techniques discrete wavelet transform and probabilistic neural network. The author in [8] developed an ANN model using a back-propagation learning rule and achieved higher classification accuracies for both the training (95%) and testing (82%) data sets. In [9], the author addresses the problem of classifying eight different heartbeat conditions using arrhythmia classification method and achieved 92.74% classification rate, whereas in [10] the author made use of image data and classified human gender and age using local gabor binary pattern histogram and yield an accuracy of 94.17%. In [11], the author uses multi-crop polling operation for feature selection on image data in identifying the proper lung functioning. In [12], the author built their own ANN to classify pollutants made up of sixty-one input neuron, three output neurons with three hidden layers. In [13], the author makes use of deep convolution neural network and support vector machine in classifying EEG signal images and achieved the average accuracy 92.24%. In [14], the author combined convolution neural network with extreme learning machine for hyperspectral imagery without spectral distortion. In [15], the author made an investigated on multi-distribution deep neural network in intonation classification of English language and yield the classification accuracy rate of 93.0%. In [16], the author modified version of the particle swarm optimization algorithm is used to train the radial basis function network for classification of the electroencephalogram signal for epileptic seizure identification, proposed method produced a maximum accuracy of 99%. In [17], the author set up an artificial neural networks model using standard merceological parameters to classify NIR data. In [18], author uses 47 images of teeth and uses wavelet Fourier descriptor in extracting features of teeth with support vector machine, achieved classification accuracy of 94% and repeated the same for about 30 CT cases.

3 Methodology

Let us consider, the input data collection as $X = \{I_1, I_2, \ldots, I_n\}$ and its associated set of label vectors $Y = \{L_1, L_2, \ldots, L_n\}$. For a two-dimensional classification problem, a possible GONN classifier is represented by a single GONN tree (GONN T (x)). For any simple pattern assume x, the single GONN tree can be constructed for two classes as in Eq. 1:

$$\left\{ \begin{array}{ll} \text{class 1} & \text{GONN}(X) \geq 0 \\ \text{class 2} & \text{GONN}(X) < 0 \end{array} \right\} \tag{1}$$

By using relative mean squared error (MSE) of the GONN, one can easily evaluate the fitness value of proposed algorithm as given in Eq. 2.

$$F_c = \frac{1}{1 + \frac{1}{x}\sum_y^x (D_y - A_y)} \tag{2}$$

where X is the length of training data set, D_y is the expected output on y and A_y is the GONN expected output, when Y is given as input.

3.1 Identifying a Collection of Input Parameters and Corresponding Collection of Output Parameters

We can start neural network toolbox in MATLAB with nnstart command. The GUI of ANN tool can be displayed. The screen wizard to select input data set and target data set. The target data is prepared based on the class labels. The target data can have many number of vectors. Each vector should consist of either 0 or 1. We have selected 15% of total input data set samples for validation, another 15% for testing. Remaining 70% input data set will be for training. Training is the process to adjust the network weights and bias according to its error. NN toolbox provides wide variety of training algorithms, each having their own learning inbuilt function. Validation is the process to measure the process generalization and to halt the training when generalization stops improving. Testing is the process to measure the performance of network during and after training. This will never affect on training.

To extract the wanted columns from the data set, the below given code is used in order to generate required output file (.txt) comma as separator. The generated output file has only the classified field called class obtaining values {0, 1}.

3.2 Structure of ANN

The hidden layer parameter is set to 10. Hidden layer: it is a neuron in feedforward network that connects as input to another layer. The hidden layer processes the weights and biases using sigmoid function.

Output layer: it is another neuron that will take the input from hidden layer and process with softmax function.

Output: it is set to 1. The out neuron is defined itself based on the target vector.

3.3 Training the System with the Help of MATLAB Software

To show the wizard of training, the network using scaled conjugate gradient back-propagation process. The performance of network will show based on cross-entropy and confusion matrices. Representing the algorithm details and numbers of iteration made to train the entire system.

4 Results

Figure 1 represents the best validation performance which occurrs at epoch 17. Figure 2 shows the confusion matrices of network. The bottom matrices will define the accuracy of the network. Here we got 73:7% for classified correctly and 26:3% are misclassified. If the network result is not satisfactory, then network can retrain by adjusting network size. Cross-entropy is to be evaluated for each class of output variables as in Eq. 3.

$$\text{Entropy} = -\,\text{training\_rate} \times \log(\text{Output}) \tag{3}$$

The aggregate cross-entropy performance is the mean of the individual values as in Eq. 4:

$$\text{Performance of Model} = \frac{\text{sum(Entropy Values)}}{\text{nume1 (Entropy)}} \tag{4}$$

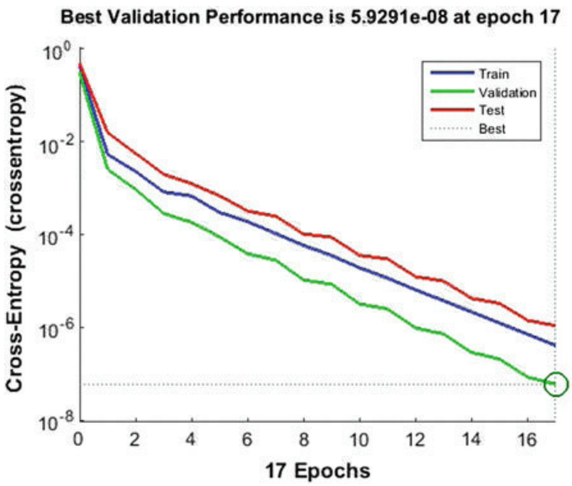

Fig. 1 Best validation performance

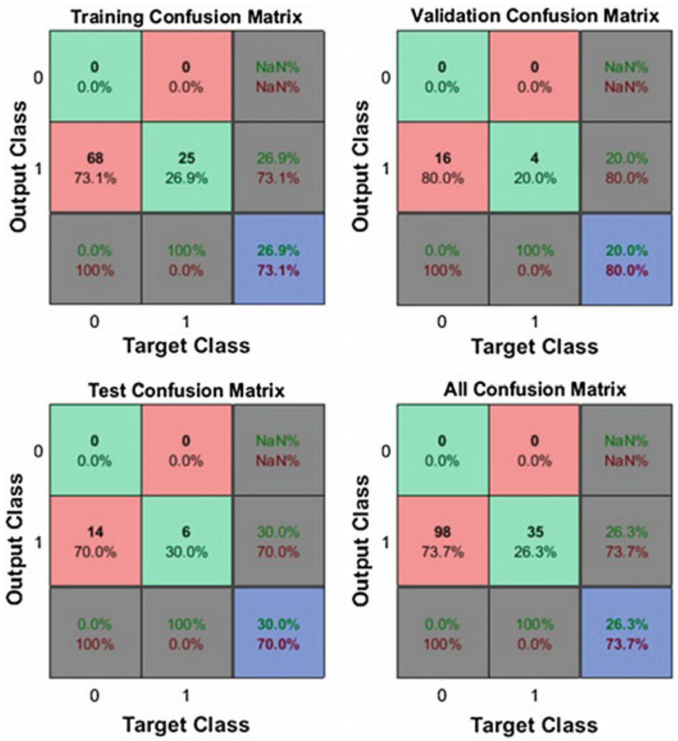

Fig. 2 Confusion Matrix