Signals and Communication Technology

This series is devoted to fundamentals and applications of modern methods of signal processing and cutting-edge communication technologies. The main topics are information and signal theory, acoustical signal processing, image processing and multimedia systems, mobile and wireless communications, and computer and communication networks. Volumes in the series address researchers in academia and industrial R&D departments. The series is application-oriented. The level of presentation of each individual volume, however, depends on the subject and can range from practical to scientific.

Indexing: All books in "Signals and Communication Technology" are indexed by Scopus and zbMATH

More information about this series at http://www.springer.com/series/4748

Andrea Abrardo · Mauro Barni ·
Kassem Kallas · Benedetta Tondi

Information Fusion in Distributed Sensor Networks with Byzantines

Springer

Andrea Abrardo
Department of Information Engineering and Mathematics
University of Siena
Siena, Italy

Mauro Barni
Department of Information Engineering and Mathematics
University of Siena
Siena, Italy

Kassem Kallas
Department of Information Engineering and Mathematics
University of Siena
Siena, Italy

Benedetta Tondi
Department of Information Engineering and Mathematics
University of Siena
Siena, Italy

ISSN 1860-4862 ISSN 1860-4870 (electronic)
Signals and Communication Technology
ISBN 978-981-32-9000-6 ISBN 978-981-32-9001-3 (eBook)
https://doi.org/10.1007/978-981-32-9001-3

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

About This Book

Every day, we share our personal information through digital systems which are constantly exposed to threats. For this reason, security-oriented disciplines of signal processing have received increasing attention in the last decades: multimedia forensics, digital watermarking, biometrics, network monitoring, steganography and steganalysis are just a few examples. Even though each of these fields has its own peculiarities, they all have to deal with a common problem: the presence of one or more adversaries aiming at making the system fail. Adversarial Signal Processing lays the basis of a general theory that takes into account the impact that the presence of an adversary has on the design of effective signal processing tools.

By focusing on the application side of Adversarial Signal Processing, namely adversarial information fusion in distributed sensor networks, and adopting a game-theoretic approach, this book presents the recent advances in the field and how several issues have been addressed. First, a heuristic decision fusion setup is presented together with the corresponding soft isolation defense scheme that protects the network from adversaries, specifically, Byzantines. Second, the development of an optimum decision fusion strategy in the presence of Byzantines is outlined. Finally, a technique to reduce the complexity of the optimum fusion by relying on a novel nearly optimum message passing algorithm based on factor graphs is presented.

Contents

1 Introduction 1
1.1 Motivation 1
1.2 Goal and Summary 5
1.2.1 Goal 5
1.2.2 Summary of the Book 6
References 7

2 Basic Notions of Distributed Detection, Information Fusion and Game Theory 9
2.1 Introduction 9
2.2 Detection Theory 10
2.2.1 Bayesian Detection 11
2.2.2 Detection Performance Metrics 13
2.2.3 Neyman-Pearson Detection 14
2.2.4 Sequential Detection 15
2.3 Information Fusion Rules 16
2.3.1 Simple Fusion Rules 17
2.3.2 Advanced Fusion Rules 19
2.4 Game Theory in a Nutshell 21
2.4.1 Nash Equilibirium 23
2.4.2 Dominance Solvable Games 24
2.5 Conclusion 26
References 26

3 Security Attacks and Defenses in Distributed Sensor Networks 29
3.1 Introduction 29
3.2 Attacks to Distributed Sensor Networks 29
3.2.1 Attacks to the Observations 31

3.2.2 Attacks to the Sensors 33
3.2.3 Attacks to the Reports 34
3.3 Defenses Against Attacks to Distributed Sensor Networks 34
3.3.1 Defenses Against Attacks to the Observations 34
3.3.2 Defenses Against Attacks to Sensors 37
3.3.3 Defenses Against Attacks to Reports 40
3.4 Conclusion 40
References 41

4 Adversarial Decision Fusion: A Heuristic Approach 45
4.1 Introduction 45
4.2 Decision Fusion with Isolation of Byzantines 46
4.2.1 Problem Formulation 46
4.2.2 Byzantine Identification: Hard Reputation Measure 48
4.3 Decision Fusion with Soft Identification of Malicious Nodes 48
4.4 A Game-Theoretical Approach to the Decision Fusion Problem 50
4.4.1 The Decision Fusion Game 50
4.4.2 Equilibrium Point Analysis of the Decision Fusion Game 51
4.5 Performance Analysis 52
4.6 Conclusions 55
References 55

5 A Game-Theoretic Framework for Optimum Decision Fusion in the Presence of Byzantines 57
5.1 Introduction 57
5.2 Optimum Fusion Rule 58
5.2.1 Unconstrained Maximum Entropy Distribution 61
5.2.2 Constrained Maximum Entropy Distributions 61
5.2.3 Fixed Number of Byzantines 64
5.3 An Efficient Implementation Based on Dynamic Programming 64
5.4 Optimum Decision Fusion Game 66
5.5 Simulation Results and Discussion 68
5.5.1 Equilibrium Point of the DF_{Byz} game 68
5.5.2 Performance at the Equilibrium 76
5.5.3 Assumptions Validation and Discussion 78
5.6 Conclusions 80
References 80

6 An Efficient Nearly-Optimum Decision Fusion Technique Based on Message Passing . . . 83
6.1 Introduction . . . 83
6.2 Notation and Problem Formulation . . . 84
6.3 A Decision Fusion Algorithm Based on Message Passing . . . 86
6.3.1 Introduction to Sum-Product Message Passing . . . 86
6.3.2 Nearly-Optimal Data Fusion by Means of Message Passing . . . 88
6.4 Simulation Results and Discussion . . . 93
6.4.1 Complexity Assessment . . . 94
6.4.2 Performance Evaluation . . . 94
6.5 Conclusions . . . 100
References . . . 101

7 Conclusion . . . 103
7.1 Open Issues . . . 103
Reference . . . 104

Bibliography . . . 105

Index . . . 107

Symbols

H_0	Null hypothesis
H_1	Alternative hypothesis
n	Number of nodes in the network
$\mathbf{x}_i$	Observation vectors available to sensor i
S_i	The system state under hypothesis $H_i, i \in \{0, 1\}$
$P(H_0)$	A-priori probability that the system is in state S_0
$P(H_1)$	A-priori probability that the system is in state S_1
$P(x\|H_j)$	The observation probability density conditioned to hypothesis H_j
$S^* \in \{0, 1\}$	The global decision at the fusion center regarding S
C_{ij}	Cost of deciding H_i when H_j is true
$\mathcal{C}$	Average cost or risk function for Bayesian detection
$\Lambda(x)$	Likelihood ratio regarding the observation x
λ	Decision threshold
P_{fa}	Probability of false alarm
P_{md}	Probability of missed detection
P_d	Probability of correct detection
P_{null}	Probability to decide H_0 when H_0 is true
P_e	Probability of error
λ_{NP}	Local Neyman-Pearson likelihood decision threshold
α_{NP}	Acceptable false alarm for Neyman-Pearson detector
$\mathcal{F}$	Lagrange function for Neyman-Pearson detector optimization
$\lambda_i, i \in \{0, 1\}$	Decision threshold for hypothesis H_i for local SPRT detector
α_{ST}	Local SPRT detector constraint on false alarm probability
β_{ST}	Local SPRT detector constraint on missed detection probability
u_i	Information sent by sensor i to the FC
P_{d_i}	Local probability of correct detection at node i
P_{fa_i}	Local probability of false alarm at node i
P_{md_i}	Local probability of missed detection at node i
Q_D	Global probability of correct detection at the FC

Q_{FA}	Global probability of false alarm at the FC
$Q_{D_{AND}}$	Global probability of correct detection for the AND rule
$Q_{FA_{AND}}$	Global probability of false alarm for the AND rule
$Q_{D_{OR}}$	Global probability of correct detection for the OR rule
$Q_{FA_{OR}}$	Global probability of false alarm for the OR rule
$Q_{D_{kn}}$	Global probability of correct detection for the k-out-of-n rule
$Q_{FA_{kn}}$	Global probability of false alarm for the k-out-of-n rule
U_{SLC}	Square Law Combining information fusion result
U_{MRC}	Maximum Ratio Combining information fusion result
U_{SC}	Selection Combining information fusion result
ζ	Decision threshold of the soft combination rules
$\Upsilon_i, i \in \{0, 1\}$	Decision threshold for hypothesis H_i for global SPRT detector
α_{FC}	Global SPRT detector constraint on false alarm probability
β_{FC}	Global SPRT detector constraint on missed detection probability
$\mathcal{X}_M^2$	Chi-square distribution with M degrees of freedom
$\Gamma(.)$	The incomplete gamma function
$Q(.)$	The generalized Marcum *Q-function*
$\mathcal{S}_i$	Strategy set available to player i
v_l	Payoff (or utility) of player l
$G(N, \mathcal{S}, \mathbf{v})$	Game with N players, strategy set $\mathcal{S}$ and payoff vector $\mathbf{v}$
$\Pi(\mathcal{Z})$	Set of all probability distributions over the set $\mathcal{Z}$
r_i	Report sent by node i to the FC
α	Fraction of nodes (or links) under attack or the probability that a node (or link) is under attack
r_{ij}	Report by node i at instant j
m	Observation window size
P_{mal}	Node malicious probability or crossover probability of the attacked links
u_{ij}	Decision by node i at instant j
Γ_i	Hard reputation score of node i
$d_{int}(j)$	Intermediate decision at instant j at the FC
η	Isolation threshold
R_{ij}	Soft reputation score of node i at instant j
$DF(\mathcal{S}_{FC}, \mathcal{S}_{FC}, v)$	Decision fusion game with $\mathcal{S}_{FC}$ the strategy set for the FC, $\mathcal{S}_B$ the strategy set for Byzantines, and payoff v
$P_{e,ar}$	Probability of error after removal of Byzantines
P_{ISO}^B	Probability of correct isolation of Byzantines
P_{ISO}^H	Probability of erroneous isolation of honest nodes
P_{mal}^{FC}	The FC guess of P_{mal}
$P_X(x)$	Probability mass function of the random variable x
S^m	Sequence of system states random variable with instantiation s^m
$P_{S_j}(i), i \in \{0, 1\}$	Probability that a system is in state S_j at time i

U_{ij}	Random variable for the local decision of node i at instant j with instantiation u_{ij}
$A^n = (A_1, \ldots, A_n)$	Binary random sequence for Byzantine positions with a^n, its instantiation
$\mathbf{R} = \{R_{ij}\}$	Random matrix of all received reports by FC with $\mathbf{R} = \{R_{ij}\}$ as its instantiation
$P(a^n)$	Probability of Byzantine sequence
ε	Local decision error at the nodes
δ	The probability that the FC receives a wrong report
$m_{eq}^{(i)}$	The number of instants at which the report is equal to the system state for node i
$E[N_B]$	Expected number of Byzantines
μ_{A_i}	Expected value of A_i
$H(A^n)$	Entropy distribution of Byzantines
$h(\mu_{A_i})$	Binary entropy function for the expected value of A_i
h	The FC expected maximum number of Byzantines
$\mathcal{I} = \{1, \ldots, n\}$	Indexing set of size n
$\mathcal{I}_k$	Set of all k-subsets of $\mathcal{I}$
I	Random variable with indexes of Byzantine nodes
$P(I)$	Equivalent to the probability of a Byzantine sequence $P(a^n)$
n_B	Fixed number of Byzantines in the network known to the FC
$DF_{Byz}(\mathcal{S}_B, \mathcal{S}_{FC}, v)$	Decision fusion game with $\mathcal{S}_B$ the strategy set of Byzantines, $\mathcal{S}_{FC}$ the strategy set of the FC, and v the payoff
P_{mal}^B	Malicious probability strategy of the Byzantines
$\mathcal{S}_B^q$	Quantized Byzantines' strategy set
$\mathcal{S}_{FC}^q$	Quantized FC's strategy set with $\mathbf{r} = \{r_{ij}\}$ as its instantiation
$\mathbf{V}$	Payoff matrix for each pair of strategies
P_e^*	Probability of error at the equilibrium
$P(P_{mal}^B)$	Probability assigned by Byzantines to a strategy in mixed strategy Nash equilibrium
$P(P_{mal}^{FC})$	Probability assigned by FC to a strategy in mixed strategy Nash equilibrium
ρ	State transition probability in a two-state Markov model
$m_{vf^{(l)}}$	Variable-to-function message for factor l
$m_{fv}^{(l)}$	Function-to-variable message for factor l

List of Figures

Fig. 2.1 Parallel Topology 11
Fig. 2.2 ROC curve example 14
Fig. 2.3 Neyman-Pearson Setup 15
Fig. 2.4 SPRT detector. 16
Fig. 3.1 Classification of attacks to distributed sensor networks. 30
Fig. 4.1 Decision fusion under adversarial conditions 46
Fig. 4.2 Error probability $P_{e,ar}$ at the equilibrium for $P_d = 0.8$ (**a**) and $P_d = 0.9$ (**b**). 54
Fig. 4.3 P^H_{iso} versus P^B_{iso} at $P_{\text{mal}} = 1.0$, for $\alpha = 0.46$ and $P_d = 0.8$. For the soft scheme, 10 thresholds are taken 55
Fig. 5.1 Sketch of the adversarial decision fusion scheme 58
Fig. 5.2 Efficient implementation of the function in (5.18) based on dynamic programming. The figure depicts the tree with the iterations for the case $k < n - k$ 66
Fig. 6.1 Markovian model for system states. When $\rho = 0.5$ subsequent states are independent. 85
Fig. 6.2 Node-to-factor message passing 88
Fig. 6.3 Factor-to-node message passing 88
Fig. 6.4 End of message passing for node z_i 89
Fig. 6.5 Factor graph for the problem in Eq. (6.10). 89
Fig. 6.6 Factor graph for the problem at hand with the illustration of all the exchanged messages 91
Fig. 6.7 Number of operations required for different n, $m = 10$ and 5 message passing local iterations for message passing and optimal schemes. 95
Fig. 6.8 Number of operations required for different m, $n = 20$ and 5 message passing local iterations for message passing and optimal schemes. 95

Fig. 6.9 Error probability as a function of α for the following setting: $n = 20$, independent Sequence of States $\rho = 0.5$, $\varepsilon = 0.15$, $m = 10$ and $P_{\mathrm{mal}} = 1.0$ 96
Fig. 6.10 Error probability as a function of α for the following setting: $n = 20$, Markovian Sequence of States $\rho = 0.95$, $\varepsilon = 0.15$, $m = 10$ and $P_{\mathrm{mal}} = 1.0$ 97
Fig. 6.11 Error probability as a function of α for the following setting: $n = 20$, Markovian Sequence of States $\rho = 0.95$, $\varepsilon = 0.15$, $m = 30$ and $P_{\mathrm{mal}} = 1.0$ 98
Fig. 6.12 Error probability as a function of α for the following setting: $n = 20$, Markovian Sequence of States $\rho = 0.95$, $\varepsilon = 0.15$, $m = 30$ and $P_{\mathrm{mal}} = 0.5$ 98
Fig. 6.13 Error probability as a function of m for the following settings: $n = 20$, Markovian Sequence of States $\rho = 0.95$, $\varepsilon = 0.15$ and $\alpha = 0.45$. 99
Fig. 6.14 Error probability as a function of m for the following settings: $n = 20$, independent Sequence of States $\rho = 0.5$, $\varepsilon = 0.15$ and $\alpha = 0.45$. 99
Fig. 6.15 Comparison between the case of independent and Markovian system states ($n = 20$, $\rho = \{0.5, 0.95\}$, $\varepsilon = 0.15$, $m = 10$, $P_{\mathrm{mal}} = 1.0$) 100

List of Tables

Table 2.1 Decision cases in binary detection 12

Table 2.2 Example of game representation in normal form. The row player is player 1 and the column player is player 2. The entries of the table are the payoffs of the game for each pair of strategies. 22

Table 2.3 Example of removal of weakly dominated strategies will cause the loss of some Nash equilibria. The row player is player 1 and the column player is player 2 25

Table 4.1 Payoff of the DF_H game for $\alpha = 0.46$ and $P_d = 0.8$, $P_{fa} = 0.2$ 53

Table 4.2 Payoff of the DF_S game for $\alpha = 0.46$ and $P_d = 0.8$, $P_{fa} = 0.2$ 53

Table 5.1 Payoff of the DF_{Byz} game ($10^3 \times P_e$) with independent node states with $\alpha = 0.3$, $m = 4$, $n = 20$, $\varepsilon = 0.1$. The equilibrium point is highlighted in bold 71

Table 5.2 Payoff of the DF_{Byz} game ($10^2 \times P_e$) with independent node states with $\alpha = 0.4$, $m = 4$, $n = 20$, $\varepsilon = 0.1$. The equilibrium point is highlighted in bold 71

Table 5.3 Payoff of the DF_{Byz} game ($10^2 \times P_e$) with independent node states with $\alpha = 0.45$, $m = 4$, $n = 20$, $\varepsilon = 0.1$. The equilibrium point is highlighted in bold 71

Table 5.4 Payoff of the DF_{Byz} game ($10^4 \times P_e$) with $n_B = 6$, $m = 4$, $n = 20$, $\varepsilon = 0.1$. The equilibrium point is highlighted in bold 72

Table 5.5 Payoff of the DF_{Byz} game ($10^3 \times P_e$) with $n_B = 8$, $m = 4$, $n = 20$, $\varepsilon = 0.1$. No pure strategy equilibrium exists 72

Table 5.6 Payoff of the DF_{Byz} game ($10^2 \times P_e$) with $n_B = 9$, $m = 4$, $n = 20$, $\varepsilon = 0.1$. The equilibrium point is highlighted in bold 72

Table 5.7 Mixed strategies equilibrium for the DF_{Byz} game with $n_B = 8$, $m = 4$, $n = 20$, $\varepsilon = 0.1$. P_e^* indicates the error probability at the equilibrium . . . 72
Table 5.8 Payoff of the DF_{Byz} game ($10^2 \times P_e$) with $N_B < n/2$. The other parameters of the game are set as follows: $m = 4$, $n = 20$, $\varepsilon = 0.1$. The equilibrium point is highlighted in bold . . . 73
Table 5.9 Payoff of the DF_{Byz} game ($10^4 \times P_e$) with $N_B < n/3$. The other parameters of the game are set as follows: $m = 4$, $n = 20$, $\varepsilon = 0.1$. The equilibrium point is highlighted in bold . . . 73
Table 5.10 Payoff of the DF_{Byz} game ($10^3 \times P_e$) with independent node states with $\alpha = 0.3$, $m = 10$, $n = 20$, $\varepsilon = 0.1$. The equilibrium point is highlighted in bold . . . 74
Table 5.11 Payoff of the DF_{Byz} game ($10^2 \times P_e$) with independent node states with $m = 10$, $n = 20$, $\alpha = 0.4$, $\varepsilon = 0.1$. The equilibrium point is highlighted in bold . . . 74
Table 5.12 Payoff of the DF_{Byz} game ($10^2 \times P_e$) with independent node states with $\alpha = 0.45$, $m = 10$, $n = 20$, $\varepsilon = 0.1$. The equilibrium point is highlighted in bold . . . 75
Table 5.13 Payoff of the DF_{Byz} game ($10^4 \times P_e$) with $n_B = 6$, $m = 10$, $n = 20$, $\varepsilon = 0.1$. The equilibrium point is highlighted in bold . . . 75
Table 5.14 Payoff of the DF_{Byz} game ($10^4 \times P_e$) with $n_B = 8$, $m = 10$, $n = 20$, $\varepsilon = 0.1$. No pure strategy equilibrium exists . . . 75
Table 5.15 Payoff of the DF_{Byz} game ($10^4 \times P_e$) with $n_B = 9$, $m = 10$, $n = 20$, $\varepsilon = 0.1$. No pure strategy equilibrium exists . . . 75
Table 5.16 Payoff of the DF_{Byz} game ($10^4 \times P_e$) with $N_B < n/2$. The other parameters of the game are set as follows: $m = 10$, $n = 20$, $\varepsilon = 0.1$. No pure strategy equilibrium exists . . . 76
Table 5.17 Payoff of the DF_{Byz} game ($10^4 \times P_e$) with $N_B < n/3$ in the following setup: $m = 10$, $n = 20$, $\varepsilon = 0.1$.The equilibrium point is highlighted in bold . . . 76
Table 5.18 Mixed strategies equilibrium for the DF_{Byz} game with $n_B = 8$, $m = 10$, $n = 20$, $\varepsilon = 0.1$. P_e^* indicates the error probability at the equilibrium . . . 76
Table 5.19 Mixed strategies equilibrium for the DF_{Byz} game with $n_B = 9$, $m = 10$, $n = 20$, $\varepsilon = 0.1$. P_e^* indicates the error probability at the equilibrium . . . 76
Table 5.20 Mixed strategies equilibrium for the DF_{Byz} game with $N_B < n/2$ with $m = 10$, $n = 20$, $\varepsilon = 0.1$. P_e^* indicates the error probability at the equilibrium . . . 77

Table 5.21 Error probability at the equilibrium for various fusion schemes. All the results have been obtained by letting $m = 4$, $n = 20$, $\varepsilon = 0.1$ 77
Table 5.22 Error probability at the equilibrium for various fusion schemes. All the results have been obtained by letting $m = 10$, $n = 20$, $\varepsilon = 0.1$ 78
Table 5.23 Payoff of the DF_{Byz} game with independent node states with $\alpha_{FC} = 0.2$, $\alpha = 0.3$, $m = 4$, $n = 20$, $\varepsilon = 0.1$. The equilibrium point is highlighted in bold 79
Table 5.24 Payoff of the DF_{Byz} game with independent node states with $\alpha_{FC} = 0.2$, $\alpha = 0.4$, $m = 4$, $n = 20$, $\varepsilon = 0.1$. The equilibrium point is highlighted in bold 79
Table 5.25 Payoff of the DF_{Byz} game with $N_{B_{FC}} < n/4$ in the following setup: $m = 4$, $n = 20$, $\varepsilon = 0.1$, $N_B < n/2$. The equilibrium point is highlighted in bold 79
Table 5.26 Payoff of the DF_{Byz} game with $N_{B_{FC}} < n/6$ in the following setup: $m = 4$, $n = 20$, $\varepsilon = 0.1$, $N_B < n/2$. The equilibrium point is highlighted in bold 80
Table 6.1 Running Time (in seconds) for the Optimal and the Message Passing Algorithms for: $m = 10$, $\varepsilon = 0.15$, Number of Trials $= 10^5$ and Message Passing Iterations $= 5$. 96