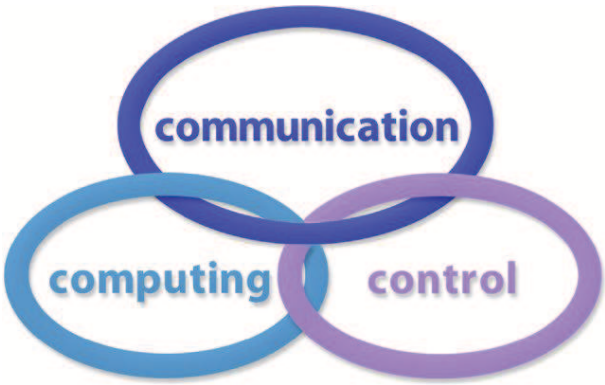
INTERNATIONAL JOURNAL

of

COMPUTERS COMMUNICATIONS & CONTROL

ISSN 1841-9836, e-ISSN 1841-9844



A Bimonthly Journal
With Emphasis on the Integration of Three Technologies

Year: 2019 Volume: 14 Issue: 3 Month: June

This journal is a member of, and subscribes to the principles of, the Committee on Publication Ethics (COPE).

http://univagora.ro/jour/index.php/ijccc/

CCC Publications

BRIEF DESCRIPTION OF JOURNAL

Publication Name: International Journal of Computers Communications & Control.
Acronym: IJCCC; **Starting year of IJCCC:** 2006.
ISO: Int. J. Comput. Commun. Control; **JCR Abbrev:** INT J COMPUT COMMUN.
International Standard Serial Number: ISSN 1841-9836, e-ISSN 1841-9844.
Publisher: CCC Publications - Agora University of Oradea.
Publication frequency: Bimonthly: Issue 1 (February); Issue 2 (April); Issue 3 (June); Issue 4 (August); Issue 5 (October); Issue 6 (December).
Founders of IJCCC: Ioan DZITAC, Florin Gheorghe FILIP and Misu-Jan MANOLESCU.

Indexing/Coverage:

- Since 2006, Vol. 1 (S), IJCCC is covered by Clarivate Analytics and is indexed in ISI Web of Science/Knowledge: Science Citation Index Expanded.

  2018 Journal Citation Reports® Science Edition (Clarivate Analytics, 2017):
  Subject Category: (1) Automation & Control Systems: Q4(2009, 2011, 2012, 2013, 2014, 2015), **Q3(2010, 2016, 2017)**; (2) Computer Science, Information Systems: Q4(2009, 2010, 2011, 2012, 2015), **Q3(2013, 2014, 2016, 2017)**.

  Impact Factor/3 years in JCR: 0.373(2009), 0.650 (2010), 0.438(2011); 0.441(2012), 0.694(2013), 0.746(2014), 0.627(2015), 1.374(2016), **1.29 (2017)**.
  Impact Factor/5 years in JCR: 0.436(2012), 0.622(2013), 0.739(2014), 0.635(2015), 1.193(2016), **1.179(2017)**.

- Since 2008 IJCCC is indexed by Scopus: **CiteScore2018 = 1.56**.
  Subject Category:
  (1) Computational Theory and Mathematics: Q4(2009, 2010, 2012, 2015), **Q3(2011, 2013, 2014, 2016, 2017, 2018)**;
  (2) Computer Networks and Communications: Q4(2009), Q3(2010, 2012, 2013, 2015), **Q2(2011, 2014, 2016, 2017, 2018)**;
  (3) Computer Science Applications: Q4(2009), **Q3(2010, 2011, 2012, 2013, 2014, 2015, 2016, 2017, 2018)**.

  SJR: 0.178(2009), 0.339(2010), 0.369(2011), 0.292(2012), 0.378(2013), 0.420(2014), 0.263(2015), 0.319(2016), 0.326 (2017), **0.37 (2018)**.

- Since 2007, 2(1), IJCCC is indexed in EBSCO.

Focus & **Scope:** International Journal of Computers Communications & Control is directed to the international communities of scientific researchers in computers, communications and control, from the universities, research units and industry. To differentiate from other similar journals, the editorial policy of IJCCC encourages the submission of original scientific papers that focus on the integration of the 3 "C" (Computing, Communications, Control).

In particular, the following topics are expected to be addressed by authors:

(1) Integrated solutions in computer-based control and communications;

(2) Computational intelligence methods & Soft computing (with particular emphasis on fuzzy logic-based methods, computing with words, ANN, evolutionary computing, collective/swarm intelligence, membrane computing, quantum computing);

(3) Advanced decision support systems (with particular emphasis on the usage of combined solvers and/or web technologies).

EDITORIAL ADDRESS:
Agora University, Cercetare Dezvoltare Agora, Tineretului 8, 410526 Oradea, Bihor, Romania,
Tel./ Fax: +40 359101032, E-mail: ijccc@univagora.ro, rd.agora@univagora.ro
URL: http://univagora.ro/jour/index.php/ijccc/

EDITORIAL BOARD OF IJCCC (MEMBERS:

Constantin GAINDRIC
IMMAS, Republic of MOLDOVA
Kishinev, 277028, Academiei 5
gaindric@math.md

Xiao-Shan GAO
Academia Sinica, CHINA
Beijing 100080, China
xgao@mmrc.iss.ac.cn

Enrique HERRERA-VIEDMA
University of Granada, SPAIN
Av. del Hospicio, s/n, 18010 Granada
viedma@decsai.ugr.es

Kaoru HIROTA
Tokyo Institute of Tech., JAPAN
G3-49,4259 Nagatsuta
hirota@hrt.dis.titech.ac.jp

Arturas KAKLAUSKAS
VGTU, LITHUANIA
Sauletekio al. 11, LT-10223 Vilnius
arturas.kaklauskas@vgtu.lt

Gang KOU
SWUFE, CHINA
Chengdu, 611130
kougang@swufe.edu.cn

Heeseok LEE
KAIST, SOUTH KOREA
85 Hoegiro, Seoul 02455
hsl@business.kaist.ac.kr

George METAKIDES
University of Patras, GREECE
Patra 265 04, Greece
george@metakides.net

Shimon Y. NOF
Purdue University, USA
610 Purdue Mall, West Lafayette
nof@purdue.edu

Stephan OLARIU
Old Dominion University, USA
Norfolk, VA 23529-0162
olariu@cs.odu.edu

Gheorghe PĂUN
Romanian Academy, ROMANIA
IMAR, Bucharest, PO Box 1-764
gpaun@us.es

Mario de J. PEREZ JIMENEZ
University of Seville, SPAIN
Avda. Reina Mercedes s/n, 41012
marper@us.es

Radu-Emil PRECUP
Pol. Univ. of Timisoara, ROMANIA
Bd. V. Parvan 2, 300223
radu.precup@aut.upt.ro

Radu POPESCU-ZELETIN
Technical University Berlin, GERMANY
Fraunhofer Institute for Open CS
rpz@cs.tu-berlin.de

Imre J. RUDAS
Obuda University, HUNGARY
Budapest, Becsi ut 96b, 1034
rudas@bmf.hu

Yong SHI
Chinese Academy of Sciences, CHINA
Beijing 100190
yshi@gucas.ac.cn, yshi@unomaha.edu

Bogdana STANOJEVIC
Serbian Academy of SA, SERBIA
Kneza Mihaila 36, Beograd 11001
bgdnpop@mi.sanu.ac.rs

Athanasios D. STYLIADIS
University of Kavala, GREECE
65404 Kavala
styliadis@teikav.edu.gr

Gheorghe TECUCI
George Mason University, USA
University Drive 4440, Fairfax VA
tecuci@gmu.edu

Horia-Nicolai TEODORESCU
Romanian Academy, ROMANIA
Iasi Branch, Bd. Carol I 11, 700506
hteodor@etc.tuiasi.ro

Dan TUFIŞ
Romanian Academy, ROMANIA
13 Septembrie, 13, 050711 Bucharest
tufis@racai.ro

Edmundas K. ZAVADSKAS
VGTU, LITHUANIA
Sauletekio ave. 11, LT-10223 Vilnius
edmundas.zavadskas@vgtu.lt

INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL
ISSN 1841-9836, e-ISSN 1841-9844, Volume 14, Issue 3, June 2019.

Contents

INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL
ISSN 1841-9836, e-ISSN 1841-9844, 14(3), 293-310, June 2019.

Network and Traffic Design Aspects in Network-Coding-Enabled Wireless Networks

K. Alic, M. Mohorcic, A. Svigelj

Kemal Alic
Jozef Stefan International Postgraduate School
Jamova cesta 39, 1000 Ljubljana, Slovenia
kemalic@gmail.com

Mihael Mohorcic
1. Department of Communication Systems, Jozef Stefan Institute
Jamova cesta 39, 1000 Ljubljana, Slovenia
2. Jozef Stefan International Postgraduate School
Jamova cesta 39, 1000 Ljubljana, Slovenia
miha.mohorcic@ijs.si

Ales Svigelj*
1. Department of Communication Systems, Jozef Stefan Institute
Jamova cesta 39, 1000 Ljubljana, Slovenia
2. Jozef Stefan International Postgraduate School
Jamova cesta 39, 1000 Ljubljana, Slovenia
*Corresponding author: ales.svigelj@ijs.si

Abstract: Practical experience of using opportunistic network coding has already been gained in several real network deployments, indicating the influence of some of the fundamental characteristics of the network and the traffic load. However, these aspects have not been systematically investigated in the scope of the construction of efficient and robust large-scale network-coding-enabled wireless mesh networks. In this paper we focus on these aspects using an example of two opportunistic network-coding procedures: the well-known COPE and the Bearing Opportunistic Network coding (BON). In addition, the design aspects for network-coding-enabled wireless mesh networks and applications are discussed. We have shown that opportunistic network coding can improve the performance of different networks and supported applications in terms of throughput, delay and jitter, although the benefits are not significant in all the cases. Thus, the use of opportunistic network coding should be considered upfront during the wireless network design phase in order to obtain the greatest benefits.
Keywords: opportunistic network coding, network coding algorithm, wireless multihop networks, simulations; traffic design.

1 Introduction

Network coding has been successfully used to improve the performance of large, wireless, multi-hop networks. With network coding the conventional store-and-forward paradigm in relaying scenarios is replaced by a process-and-forward approach. Instead of forwarding packets in the same form as they are received, the intermediate nodes combine the received outgoing packets using an algebraic function. The broadcast nature of wireless media is embraced as all the packets are distributed, for free, to all the neighbours, i.e., the nodes that can overhear the transmission. All the nodes store all the received packets for possible packet-decoding purposes. Hence, a typical network-coding mechanism is composed of two main operations: (1) the coding of the outgoing packets on (intermediate) nodes and (2) the decoding of incoming packets upon their reception on nodes [19].

Network coding was originally proposed to increase the throughput of a single-source communication, for instance, a multicast stream [1], in order to achieve the maximum data-transfer rate in multicast networks [12]. In this paper, we are interested in opportunistic network coding, where the packets for transmission are selected in a classic manner and the coding opportunities are searched for only amongst the packets currently present in the output queue [10, 13, 19]. We further focus our interest in fixed, wireless mesh networks, such as metropolitan Wi-Fi networks [6] or Wireless Sensor Networks (WSNs) [22] with unicast traffic, where the nodes can be seen as fixed wireless access points for the end users.

The performance gains of opportunistic network coding in wireless mesh networks have already been proven in testbed deployments [8, 10]. The gains have also been analysed from the theoretical perspective [24] and compared to practical gains [25]. Opportunistic network coding has also been considered for streaming applications. Th code selection that takes into account video-stream characteristics, such as the distortion values and the play-out deadlines of the video packets, was studied in [15]. The well-known and widely adopted COPE [10] has been under the spotlight, while so-called support mechanisms such as scheduling [21], queue management [7,16,17] and routing support [11,14,20] were used to bring about additional gains. Most of these support mechanisms can also be used with other opportunistic network-coding approaches, such as [3, 8, 19].

In this paper we are not optimising the parameters of any particular network-coding algorithm or one of the support mechanisms, rather we are investigating the impact of different network characteristics on the performance of network-coding procedures. Thus, our interest is in traffic distributions and patterns, different network topologies and node properties. As these parameters depend on the design and the purpose of the network, it is of paramount importance for the developers of wireless mesh networks and applications to know what benefits they can expect from the use of opportunistic network coding. As a particular example, this paper deals with the design and construction of an efficient and robust, large-scale (99 nodes) network-coding-enabled, wireless mesh network that ensures the delivery of high-data-rate and low-delay services for different applications used by a very large number of distributed users. We analyse and explain the major impacts on network performance in terms of a throughput and delay analysis of the opportunistic network-coding using two different opportunistic network coding procedures, namely, the well-known COPE [10] as a representative of a procedure based on distributed information, and the recently proposed BON (Bearing Opportunistic Network Coding) [3–5], representing procedures that make coding decisions based only on local knowledge, thus reducing the overhead. Both selected network-coding procedures are positioned between the MAC and the network layer and are performing single-hop packet coding. Support mechanisms that further improve the benefits of network coding are not considered.

This article is organised as follows. In Section 2 we briefly summarize the main characteristics of the COPE and the BON. In Section 3 the methodology for a subsequent performance comparison is presented. Section 4 analyses the influence of wireless-network design parameters on the performance of opportunistic network coding from the quantity-of-service perspective (e.g., goodput). The end user perception of network coding in terms of delay and jitter is presented in Section 5. Next, coding fairness for different traffic conditions is evaluated in Section 5 before Section 6 concludes the article.

2 Considered network-coding procedures

In this study we consider COPE [10] and BON [4] for use in wireless mesh networks. Both procedures are inter-session network-coding schemes that use the XOR operation for the coding of packets. They are fully distributed, able to discover coding opportunities and can be seen as

an independent layer in the communication stack.

After receiving an encoded packet, the BON and COPE processes try to decode it. Decoding is successful if there are enough packets in the packet pool. An ACK message is sent for the native packet, but only if the recipient node is determined as one of the next hops for the packet. If there is not enough information to decode the packet, the encoded packet is dropped.

Both procedures use cumulative ACK reports, which were first presented in [10]. Cumulative ACK-report messages are broadcasted periodically, every T_u seconds. If the opportunity arises the ACK reports are attached to the regular outgoing packets. In this case cumulative ACKs can be seen as an additional header. The cumulative ACK messages reduce the overhead compared to the individual ACK messages, thus optimizing the network coding from the throughput point-of-view.

If an ACK message is not received on the sending node within a predetermined time, the re-transmission event is triggered. A native packet that has been sent out as part of the encoded packet and has not been ACKed is to be re-transmitted. In the COPE the packet is placed at the head of the output queue. In the BON the packet is placed in the re-transmission output sub-queue. This queue has a lower priority than the signalization queue, but a higher priority than the queue with regular outgoing packets. The packets in the retransmissions queue can be coded again. However, coding opportunities are only searched for within the packets in the regular output queue. This mechanism ensures that it is not possible for the same set of native packets that already failed in the decoding to be coded again.

The COPE also uses so-called reports. With these each node informs all its neighbours about which new packets it has received. The reports are broadcasted periodically, every T_u seconds. If possible, the reports are attached to regular outgoing packets or ACK messages as additional header.

At the MAC layer the BON and COPE use a pseudo-broadcast mechanism, presented in [10]. Given that the BON and COPE rely on the same mechanisms of the underlying OSI layers, in this paper we assume the use of the same MAC layer that mimics the provision of resources as in CSMA-CD. The COPE and BON coding algorithms are described in the following subsections.

2.1 COPE

In the COPE the coding process depends of the nodes' knowledge on what information (which packets) its neighbouring nodes have, through listening to the neighbours' broadcasts (i.e., regular packets or reports). Based on this knowledge the coding process is straightforward and the decoding process will have a high success rate. Information arriving through particular messages and through listening to all the broadcasted messages provides only a few coding opportunities. In the case that the information about a packet's presence at a particular neighbour's node is not available, the coding needs to make a guess regarding the situation. The guessing is made through a delivery probability that is calculated in all of the ETX-based (Expected Transmissions Count) routing protocols. The node estimates the probability that the node N has the packet p by looking at the delivery probability for the link between packet's previous hop and node N.

With all the required information, the node can code together as many packets (p_1, ..., p_n) as possible provided that none of the packets were created on the coding node, all the packets have different next hops and that there is a strong possibility that all the next hops will be able to decode the packet. The next hop can decode the packet if it has already received all but one of the packets coded together. Let the probability that a next hop has heard packet p_m be P_m. Then, the probability, P_D, that it can decode its native packet is equal to the probability that it has heard all of the $n - 1$ native packets encoded with its own, i.e.,

$$P_D = P_1 P_2 \dots P_{n-1} \tag{1}$$

The coding algorithm ensures that the decoding probability P_D for all the next hops for a given combination of encoded packets is above the threshold G.

2.2 BON

The BON-coding process [4] relies only on information gathered locally in the network-coding layer and does not record the state of the traffic flows. The BON-coding process is based on the local bearing of the packet $p_{ij}^{(R)}$ which is defined on the relay node V_R and depends on the positions of the $p_{ij}^{(R)}$ previous hop $V_i(X_i, Y_i)$ and the next hop $V_j(X_j, Y_j)$. The bearing for the packet $p_{ij}^{(R)}$ on the relay node is defined as a unit vector, calculated as:

$$\overrightarrow{b}_{ij} = \frac{(X_j - X_i, Y_j - Y_i)}{\|(X_j - X_i, Y_j - Y_i)\|} \tag{2}$$

With the BON the packets $p_{ij}^{(R)}$ and $p_{kl}^{(R)}$ are codeable on the relay node if the previous hop of $p_{ij}^{(R)}$ (V_i) is in the *vicinity* of the next hop of $p_{kl}{}^{(R)}$ (V_l) and vice versa. This is where the *vicinity* of the node V_j for the packet $p_{ij}^{(R)}$on the relay node V_R is described as an area within the shape of an infinite cone, with the apex in the packet source V_i, and the cone axis with the direction $\overrightarrow{b_{ij}}$ and the aperture $2\epsilon_{ij}^{(R)}$.

ϵ is referred to as the tolerance angle parameter and $\epsilon \geq 0$. There is one tolerance angle on each node, i.e., $\epsilon_{ij}^{(R)} = \epsilon_{kl}^{(R)} = \epsilon^{(R)}$. By increasing the parameter ϵ, the possibility of encoding multiple packets and the probability that one of the receivers will not be able to decode the packet is increased. By reducing the parameter ϵ towards zero, the coding opportunities are reduced, but the probability of a successful packet decoding on the receiving nodes is increased. In the case of $\epsilon = 0$ only two packets can be coded.

On each node the tolerance angle ϵ is adjusted based on the retransmissions-ratio parameter (RR), reflecting the success of the packet coding, which is calculated on the node V_i as:

$$RR(m)_{V_i} = \frac{K_{RmV_i}}{K_{CmV_i} K_{PmV_i}} \tag{3}$$

where $m = \{1, 2, ...\}$, $K_{CmV_i} > 0$ is the number of native packets that were sent out as part of the encoded packets, i.e., coded packets, on the node, and K_{PmV_i} is the number of packets that were passed from the network layer to the network coding layer. K_{RmV_i} is the number of network-coding-layer retransmitted packets on the node. All the values are measured in the time interval $[(m-1)T_{\epsilon U},\ m\,T_{\epsilon U})$. The obtained value is compared to the maximum threshold RR_{max} and the minimum threshold RR_{min}. Based on the comparison outcomes, the process provides a decision about increasing, decreasing or leaving the parameter ϵ unchanged.

3 Methodology for evaluating the performance

In this section we define the performance metrics and the simulation parameters for a subsequent investigation of the impact that the network and traffic characteristics have on the efficiency of the network coding.

3.1 Performance metrics

As the elementary metric reflecting the quantity of service, we observe the network goodput (g), which is the number of useful information bits delivered by the network to a certain destination per unit of time. The goodput in our graphs is defined as the sum of all the goodput on all the network nodes at a particular network load in the i-th simulation run, i.e., $g(i)$.

We further define the gain $G(i)$ in the i-th simulation run as the relative increase of goodput obtained with the network coding with respect to the goodput without the network coding:

$$G(i) = \frac{g_{NC}(i) - g_{nocoding}(i)}{g_{nocoding}(i)} 100\% \tag{4}$$

where the gain can be observed for the all the traffic in the network or just on the individual flow level. In the latter case, we refer to it as the flow gain.

As a typical Quality of Service (QoS) metric we measure the End-to-End (ETE) Delay and jitter at the application layer. Both the ETE delay and the jitter are measured separately for each particular flow, and for all the flows. Regardless of the case, we can write the following:

$$ETE(i) = \frac{\sum_{n=1}^{K_a(i)} d_n}{K_a(i)} \tag{5}$$

where d_n is the ETE delay of the n-th packet and $K_a(i)$ is the number of packets received in the application layer.

Jitter is the standard deviation from the true periodicity of a presumed periodic signal. We only measure the jitter for flows with constant inter-arrival packet rates:

$$jitter(i) = \frac{1}{F(i)} \sum_{f=1}^{F(i)} \frac{1}{K_f(i)} \sum_{n=1}^{K_f(i)} (d_{nf} - ETE_f(i))^2 \tag{6}$$

where F is the number of flows, K_f is the number of packets received in the application layer belonging to the f-th flow, and ETE_f is the ETE for the f-th flow. When we are interested in the jitter of only one particular flow, this same equation is used for the calculation and $F(i) = 1$.

The efficiency of the coding algorithms is evaluated through the number of individual overhead packets (i.e., the overhead that is attached to regular packets as headers is not taken into account in this measure).

We also consider Jain's index [9], which is primarily used to identify underutilized channels or fairness in the load distribution. We use it to identify underutilised streams.

$$J(i) = \frac{(\sum_{f=1}^{F} g_f(i))^2}{F \sum_{f=1}^{F} g_f(i)^2} \tag{7}$$

where F is the number of flows in the network and g_f is the goodput over the f-th stream.

In the results where we show the dependency of the gain on other parameters (e.g. queue length, packet length, topology, etc.), we depict the $MaxGain$, which is calculated as the average of the three highest gain values obtained among results for a given set of parameter values. We first sort the vector $G(i)$, where $i = \{1, 2, ..., I\}$ and I is the number of simulation runs from the highest to the lowest value:

$$G_s = sort\{G(1), G(2), ..., G(I)\} \tag{8}$$

and then the $MaxGain$ is:

$$MaxGain = \frac{1}{3}(G_s(1) + G_s(2) + G_s(3)) \tag{9}$$

3.2 Network-coding settings

Each investigated characteristic was evaluated in a separate scenario. To this end, two opportunistic network-coding procedures, COPE [10] and BON [4], as well as the reference scenario where network coding was not used (no coding), were considered in the evaluation. A purposely developed simulation model, presented in [2], built in the Riverbed Modeler simulation tool, was used. Each simulation run took 150 s. The results were collected between the 90th and the 150th second to observe only the steady-state conditions. Each scenario was evaluated using 30 different network loads. For each network load we made four simulation runs using different seeds.

The COPE and BON parameters remained fixed in all the scenarios, as in this study we are focusing on the impact of the network and the traffic characteristics on the efficiency of the network-coding procedures, rather than scaling (i.e., optimising) the network-coding parameters. Both the BON and COPE use up to two retransmissions in the network-coding layer. Both procedures store packets in the packet pool for 15s after the packet has been received for the first time. The BON and COPE send out cumulative ACKs at least every 0.5 s. The network-coding-layer packet re-transmissions are scheduled for 0.6 s after the initial packet transmission. If the opportunity arises, cumulative ACKs are attached to the regular outgoing packets. The COPE reports are sent out at least every 0.5 s. When possible, the report packets are also attached to the regular outgoing packets or to the cumulative ACKs. The COPE parameters were set to the values used in the test-bed presented in [10]. The BON parameters that are common to the COPE procedures are set to the same values. The other BON parameters were set to the values used in [4].

3.3 Traffic, topology and variable parameters settings

The UDP-like traffic was generated in the origin nodes. The traffic intensity was regulated through the packet inter-arrival time and the packet lengths. We used several different constant packet lengths (2, 3, 4, 5, 6, 7, 8, 9, and 10 kbits) and variable packet length between 360 and 12,000 bits. The packet size distribution had two peaks and followed the measured internet traffic, as presented in [18]. The inter-arrival time between the packets was constant or calculated using an exponential distribution. Different numbers of traffic flows were generated between selected nodes in the network. All the flows were selected in pairs (i.e., from node V_x to V_y and from nodeV_y to V_x, where each node was selected only once. We selected n node pairs from the lowest to the highest hop-count distance, thus obtaining long flows, where the packets had several opportunities to be encoded on the path. In the case of the "$all2all$" traffic-flows distribution, all the nodes generated traffic flows and sent them to all the nodes in the network, resulting in $F = N_N(N_N - 1)$ traffic flows, where N_N is the number of nodes in the observed topology. The symmetry between the two flows in the load pair was set by the traffic-symmetry ratio.

All the nodes used the same channel with a capacity of 2 Mbit/s. The packet-delivery probability between the nodes changed dynamically during the simulations and throughout each simulation run, and depended on the amount of traffic between the neighbouring nodes and the distance between the nodes. In simulation runs with high loads the packet-delivery probability was between 0.7 and 0.97.

Routing tables were calculated at the beginning of each simulation run and these were updated every 30 s. Dijkstra's algorithm was used for the route calculation, taking into account

the ETX metric.

The BON performance and the influence of the parameters on the network-coding performance were studied in different topologies to also show that the performance improvement does not depend on optimal parameter settings. We randomly deployed 99 fixed nodes over an area of $2\,km \times 2\,km$. Several node-transmission (Tx) ranges were used in order to study different network topologies with the same node distributions. The node-transmission range indicates how far the signal from the node can be detected, which is reflected in the connectivity and the wireless links between the nodes. As an example, in Fig. 1 two topologies with the shortest (280 m) and the longest (420 m) considered node-transmission ranges are depicted.

We were interested in several parameters that influence the network-coding performance. All the variable parameters, with an included reference to the section where they were used, are presented in Table 1. In general, for each case only one parameter was varied, and all the others remained fixed.

Table 1: Simulation parameters for investigated characteristic

	Transmiss. range	**Number of traffic flows**	**Traffic-symm. ratio (%)**	**Packet length (kb) Int. time dist.**	**Queue length (packets)**
Topology design (4.1)	280 - 420 m	'*all2all*'	100	10; exp	100
Traffic distribution impact (4.2)	300 m	10 - 90, '*all2all*'	0-100	10; exp	100
Packet length evaluation (4.3)	300 m	'*all2all*'	100	2 - 10, *var*; exp	100
Analysis of output queue length (4.4)	300 m	'*all2all*'	100	10; exp	10 - 200
QoS analysis (5.1)	300 m	2, 20	100	10; const.	100
Coding fairness (5.2)	300 m	20, 90	100	10; exp	100

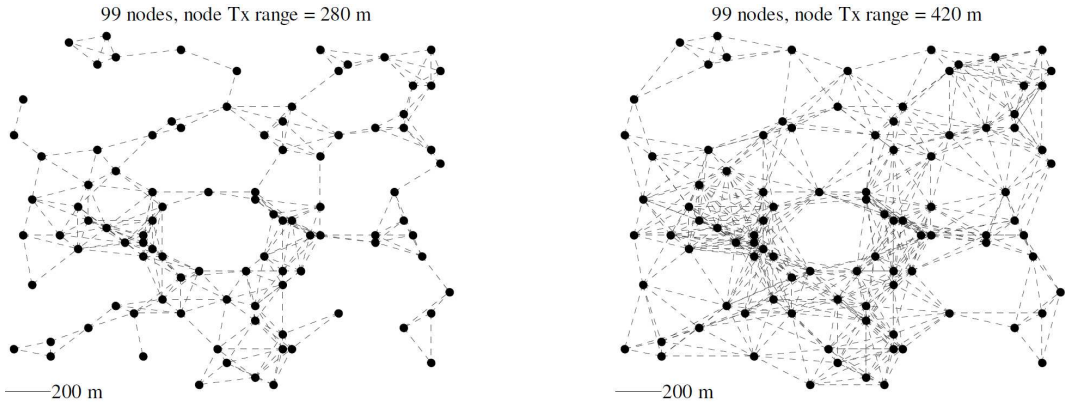


Figure 1: A graphical presentation of selected network topologies used in the evaluation

4 Quantitative impact of the network and the traffic characteristics on the network coding

The core of this section is the analysis and evaluation of the influence of the different network and traffic characteristics on the performance efficiency of the network coding in wireless mesh networks. The investigated characteristics include (i) network topology (ii) traffic distribution, (iii) packet length, and (iv) packet-queue length.

4.1 Topology design

Fig. 2 depicts the gain with respect to the increasing network load for network topologies with the node-transmission range set to 280, 340 and 400 m. The highest network-coding benefits were obtained for the topology where the nodes have the shortest transmission range. This might be counter-intuitive to a certain extent, as by a larger coverage area, it is possible to think that more nodes overhear transmissions, thus resulting in more coding opportunities and a higher decoding success. We wanted to present the gain, with its dependence on the node-transmission range, so the same results are presented in Fig. 3 as $MaxGain$, as obtained for all the topologies. The coding benefits decrease with a higher node-transmission range. There are several reasons for this. The diameter and the average hop-count distance of each network decreases with an increased nodes-transmission range. With this there are fewer possibilities for the packets to be encoded, as they need fewer hops from their sources to their destinations. In networks where the nodes have a shorter node-transmission range there are typically a few nodes that handle large portions of the traffic. Nodes that handle a larger portion of the load are where normally the majority of the coding takes place. With a larger node-transmission range the number of possible paths between the source and the destination pairs increases, hence the load is more evenly distributed among the nodes. It often happens that the flows between two neighbouring nodes on one side of the network and two neighbouring nodes on the other side of the network use entirely disjointed paths, while often still competing for the same wireless resources, but not being able to be encoded in the intermediate nodes. This can be improved, for example, by using network-coding-aware routing which could in such a case increase the coding opportunities.

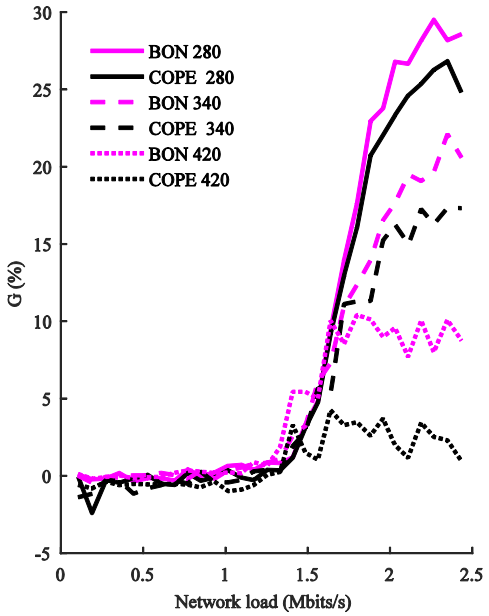


Figure 2: Dependence of the gain (G) on the network load for topologies with a 280, 340, and 400 m node-transmission range

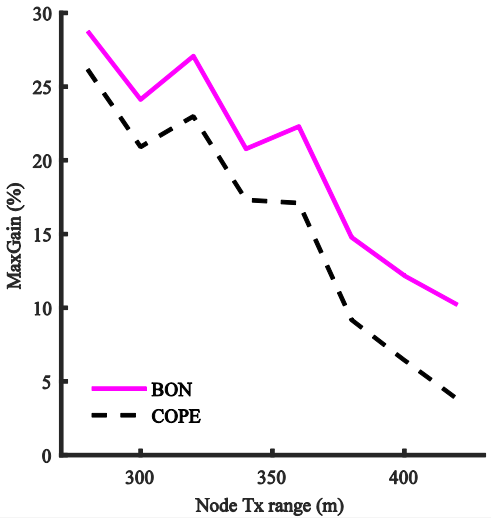


Figure 3: Dependence of the $MaxGain$ on the node-transmission (Tx) range

Furthermore, with a higher node-transmission range the competition to access the wireless channel is stiffer. Hence, the time needed to receive the network-coding-layer acknowledgement message increases, especially when the network has to deal with higher traffic loads. Thus, the acknowledgement messages can arrive at their destinations after the retransmission procedure was initiated. To some extent this can be monitored by adjusting (i.e. increasing) the waiting

time for the network-coding layer acknowledgement messages, but this might affect the streaming applications in terms of unacceptable jitter.

Based on the above analysis we can conclude that the goodput of networks where the node transmission range is high can only be increased by a few percent using basic opportunistic network-coding approaches. If the network coding is considered for use in a mesh network, where the nodes have many neighbours, the support of a network-coding-aware routing protocol should be considered, or a network-coding procedure with a different approach should be selected such as e.g. BEND [23].

4.2 Impact of the traffic distribution

The potential benefits of using network coding are strongly dependent on the traffic distribution. We first examined the influence of traffic symmetry and later on the traffic-flows distribution on the performance of opportunistic network coding.

For investigating the influence of traffic symmetry, we deployed 10 flows between 5 node pairs and varied the traffic-symmetry ratio, as defined in Table 1, where also all the other parameter settings are stated.

In Fig. 4 we present the results for the dependence of $MaxGain$ on the traffic-symmetry ratio. As expected, the gains increase as the traffic within the traffic flows becomes more symmetric, essentially increasing the network-coding opportunities. With the lowest ratio (i.e., a traffic symmetry ratio = 0) small gains were noted. Since the coding happens between the flows that cross their paths in opposite directions, the benefits of network coding increase with an increasing traffic-symmetry ratio. We can see that the network-coding procedures bring the most benefits when the traffic symmetry is above 60 %.

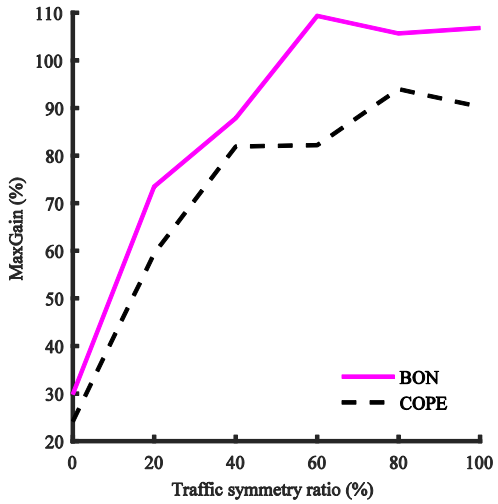


Figure 4: Dependence of the $MaxGain$ on the traffic-symmetry ratio

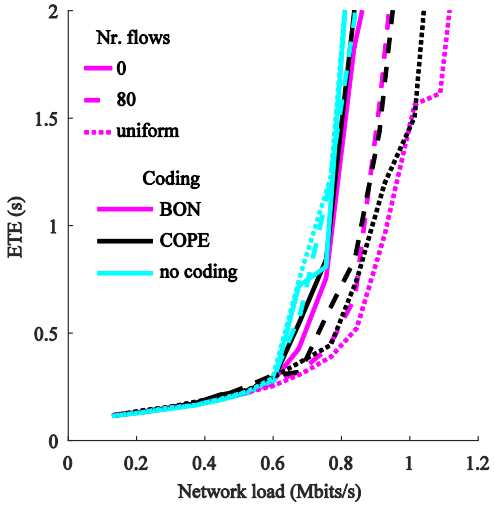


Figure 5: Dependence of the ETE delay on the goodput (g) for 20 and 90 traffic flows and '*all2all*' traffic distribution

In the next set of results, we evaluated the dependency of the network-coding benefits with respect to the number of traffic flows, i.e., the traffic-flows distribution. In Fig. 5, The dependence of the ETE delay on the goodput for 20 and 90 traffic flows and the '*all2all*' traffic distribution is presented. By increasing the number of traffic flows the network goodput increased with an unchanged ETE delay. This is expected, as with the larger number of traffic flows, the same overall network load is more evenly distributed across the network. Furthermore, the network

coding decreases the delay in comparison to the no-coding scenario, while increasing the network goodput, but the delay is increasing due to the increased network load.

In Fig. 6 the dependence of *MaxGain* on the number of traffic flows is shown. With very few flows in the network (i.e., 10) the *MaxGain* was higher than 100 %. By increasing the number of flows also shorter flows, were injected into the network, and the shorter the traffic flow, the fewer coding opportunities its packets experience along the path. The lowest gains were obtained in the '*all*2*all*' traffic-flows distribution (shown as separate points in the graph), where an in-depth analysis also revealed that the fewest coding opportunities are found. Nevertheless, in the '*all*2*all*' distribution of traffic flows, where the network-coding benefits were the lowest, the network-coding gain with both procedures was still higher than 20 %, and in addition both procedures decreased the delay significantly.

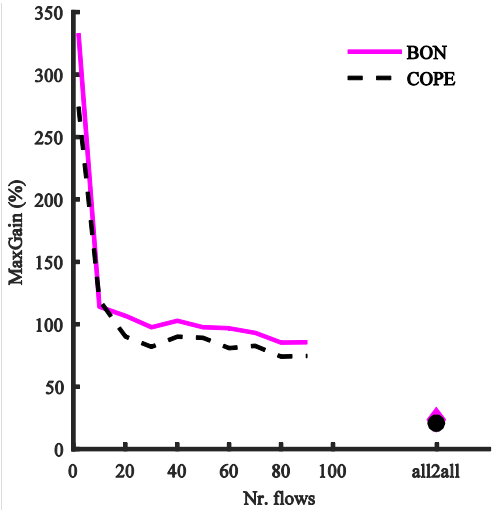


Figure 6: Dependence of the *MaxGain* on the number of traffic flows

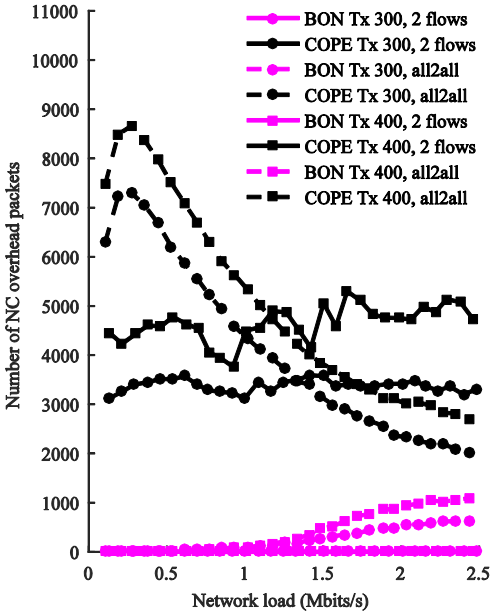


Figure 7: Network coding overhead in terms of the number of individual packets for 2 flows and '*all*2*all*' traffic distributions and topologies with node transmission (Tx) range 300 and 400 m

In topologies where the nodes have a higher node-transmission range, the signalization overhead also becomes more significant. In Fig. 7 the overhead caused by the BON and COPE in terms of the number of individual packets is presented for two extreme traffic distributions i.e. 2 flows and '*all*2*all*', and topologies with node-transmission ranges of 300 and 400 m. The COPE uses the so-called report messages, with which the nodes inform their neighbours about which packets they have received. With the increased node-transmission range there are more nodes that can overhear each other's transmissions, which means that report messages increase in number and in the size (each node composes reports for each of the neighbours from which it has received new, unreported) packets). A higher number of traffic flows means a higher number of COPE reports, as there are more nodes that can overhear the packet transmission. In low traffic-load conditions there are few opportunities to attach reports to regular outgoing packets, hence a high number of individual report messages. By increasing the traffic load the number of report messages also depends on the traffic-flow distributions. In an '*all*2*all*' traffic-flow distribution

the number of stand-alone individual messages quickly decreases, as there are plenty of regular outgoing packets on all the nodes to which the report messages can be attached, and hence fewer individual packets. The opposite happens in the case of only two traffic flows. There is a constant number of nodes that relay packets to which report messages can be attached. Due to the compact format of report messages the number of individual messages remains almost constant, regardless of the traffic load.

Using the BON, the number of overhead packets increases with the load. With higher loads there are more encoded packets and thus more acknowledgement messages which generate the only potential individual overhead packets in the BON. Similarly, as with the COPE, the acknowledgement messages are attached to regular outgoing packets, if possible. In a traffic distribution with two flows the inter-arrival time of the packets within the same flow is short enough so that the acknowledgement messages are attached to regular outgoing packets, hence there are almost no individual acknowledgement packets. It can be seen that the overhead also increases with the increased transmission range. This indicates that there can be more coding opportunities in topologies with a larger transmission range. Nevertheless, the overhead produced by the BON is significantly lower than by the COPE.

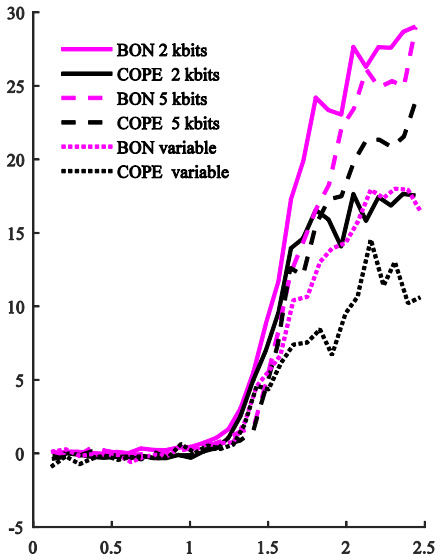


Figure 8: Dependence of the gain (G) on the network load with packets of 2000 and 5000 bits, and of variable length

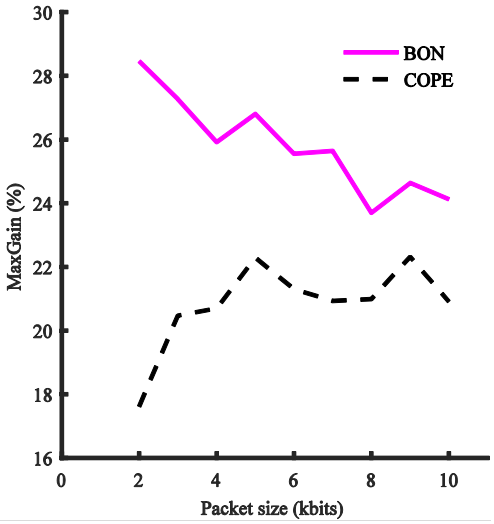


Figure 9: Dependence of the $MaxGain$ on the packet length

The influence of the packet length on the performance of opportunistic network coding reveals the overhead caused by the network-coding and how well the network coding procedures encode packets together. In this evaluation the parameters were set according to the specifications in Table 1.

In Fig. 8 the dependence of gain on the network load for packets of 2000 and 5000 bits is shown (the results for a variable packet length are also shown, but they are discussed later). We can see that the highest gains were obtained with the BON for a packet length of 2000 bits. The BON self-adaptation process is based solely on information gathered in the network-coding layer. The more packets the network-coding layer processes, the better the decision process, and hence there are fewer retransmissions, indicating better coding decisions.

However, with the BON the lowest gains are obtained with the largest packets, where, although the influence of the overhead is the lowest, incorrect coding decisions result in a high

cost of retransmission. In addition, the impact on the signalization is higher than when smaller packets are in use. The channel is occupied for longer periods and the acknowledgement-message time out can be affected, resulting in additional packet retransmissions.

The dependence of $MaxGain$ on the packet length is presented in Fig. 9 for the BON and COPE. With the BON the $MaxGain$was decreasing with the increased packet length since the overhead was less important and because the network-coding layer handled fewer packets. The latter resulted in poorer coding decisions as the BON self-adaptation process also depends on the number of processed packets. The COPE generated a higher overhead than the BON, and hence the $MaxGain$ was lower with short packets. By increasing the packet length, the gains also grew to the point where the signalization did not affect the results and the $MaxGain$ shows no dependence on the size of the packet length .

We were also interested in the network performance in terms of the variable packet lengths (also presented in Fig. 8, and noted as "*variable*"). Thus, additional experiments were carried out with the empirical packet-length distribution presented in [18], where the packet lengths varied between 360 and 12000 bits. When encoding packets with different lengths, the shorter packets are padded with zeroes, thus adding irrelevant information to the payload. As expected, the results for a variable packet-length distribution show the lowest gains. Both coding algorithms favour the coding of a packet of approximately the same length, but the degrading impact of the packet extension cannot be compensated.

4.3 Analysis of the length of the output queue

In this subsection we examine the influence of the output-queue length on the network-coding performance. The output-queue length is important because The COPE and BON do not deliberately delay packets in order to find more coding opportunities. Instead, both analysed algorithms only search for coding opportunities among the packets present in the output queue. For this analysis the parameters were set according to the specifications in Table 1.

In Fig. 10 the dependence of $MaxGain$ on the queue length, ranging from 10 to 200 packets, is shown. As expected, both network-coding procedures provided the lowest gain with the shortest queue. Although we would expect that more packets in the output queue will increase the coding opportunities, the results show that initially the $MaxGain$ grew very quickly with the queue length, but then the increase in the queue length had no effect on the performance of the network coding.

5 End-user perception of the network coding

5.1 Quality-of-service parameters

One particularly important consideration when introducing network coding in wireless communication systems is its potential effect on the quality of service. The BON and COPE use asynchronous acknowledgement mechanisms to confirm the successful transmission and decoding of encoded packets. Furthermore, cumulative acknowledgements are used to avoid a high overhead and to increase the goodput. This is reflected in the relatively long time periods used for packet-retransmission scheduling (in our case 0.6 s). Such a long waiting period is expected to affect the streaming applications in terms of the delay and jitter (i.e., QoS). We analysed this by observing a traffic flow between the two nodes with the highest hop-count distance under two traffic-flow distributions. In the considered network topology with the node-transmission range set to 300 m, these two nodes were 14 hops apart. We first investigated the case of two symmetric traffic flows between the two nodes. In such a traffic distribution the coding happened only be-

tween flows travelling in the opposite directions and retransmissions in the network coding layer were only due to packet loss and never due to an inability to decode the encoded packet. Next, we created twenty symmetric traffic flows between ten node pairs with the highest hop-count distance, but we still observed the conditions on the same node pair as before. Other parameters were set according to Table 1.

Fig. 11 shows the dependence of ETE delay on the increasing network load with two and twenty symmetric flows. The delay was not affected at lower loads in comparison to the no-coding scenario. However, with the higher loads, the delay obtained in the case of network coding was significantly lower than in the no-coding scenario. As shown in Fig. 12, the jitter also remained lower or approximately the same as in the no-coding scenario for both network-coding procedures, especially in the network with only two flows. With high loads there were occasions when the jitter increased, though this was in the case when the network was congested and packets were being dropped out of the packet queues. An in-depth analysis also revealed that with high loads there were rare occasions when the packet travelled the complete path between two nodes without getting encoded on at least one of the hops. For most of the cases on a 14-hop path the packets were encoded on several hops.

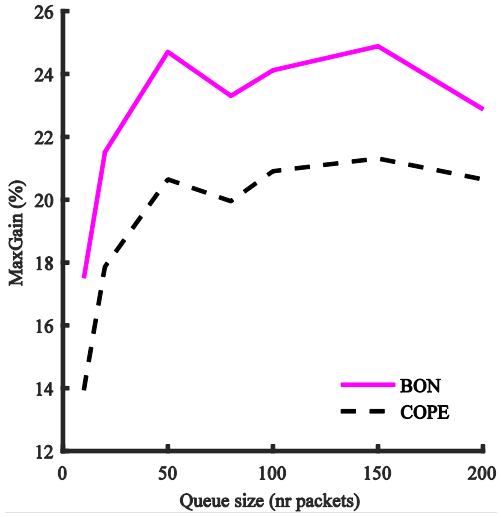


Figure 10: Dependence of the $MaxGain$ on the queue length

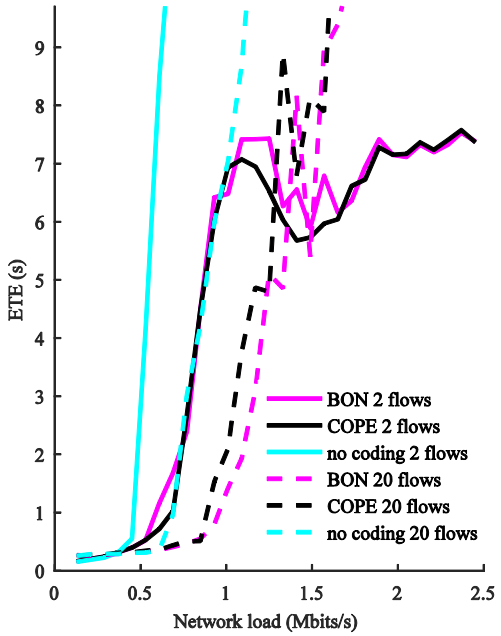


Figure 11: Dependence of the ETE delay for the longest flow on the network load with 2 and 20 symmetric traffic flows.

The jitter for a network load with 20 symmetric flows shows similar results. Under light network-load conditions the jitter obtained for both the BON and COPE was slightly higher than for the no-coding scenario. With a low number of packets in the network the impact of every encoded packet that was unsuccessfully delivered or decoded is very significant. With an increased load the jitter generally remained lower than in the no-coding scenario, although there were a few occurrences when the jitter for the case of using network coding, especially COPE, was high. This happened when a higher number of opportunistically coded packets got lost during the transmissions and the arrival order of the packets was significantly affected.

Generally, we can conclude that network coding has a positive impact on the jitter. The benefits of network coding with respect to the no-coding scenario appear at high network loads,

when the jitter is significantly affected in the no-coding scenario due to the congestion of the most-loaded nodes. It is worth pointing out that the impact of network coding was observed on a large-scale network using traffic flows that travel long distances. Such networks are not primarily built for streaming applications with high QoS demands, but were selected for the presented analysis so that packets were encoded several times on their path to the destination, and hence the effects of network coding would be observed with higher reliability.

5.2 Analysis of coding fairness

An important consideration for the network-coding procedure is also its fairness, i.e., how different traffic flows benefit from the network coding. In general traffic conditions only a small part of the native packets are encoded, thus observing the performance of the network coding for the entire network may blur the effects at the level of individual traffic flows. The aim of this section is to analyse the fairness of the network coding with respect to the traffic flows in the network, i.e., if it improves the network performance by significantly improving some traffic flows at the expense of some others. It is desirable that the network-coding algorithm improves the performance of all the traffic flows or at least that it does not degrade their performance in comparison to the no-coding scenario.

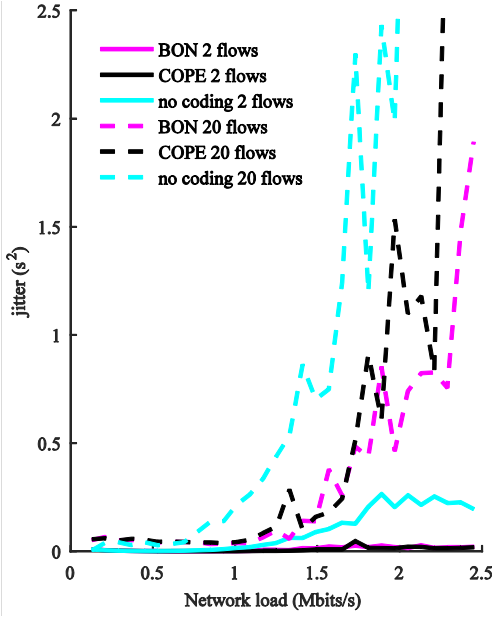


Figure 12: Dependence of the jitter for the longest flow on the network load with 2 and 20 symmetric traffic flows

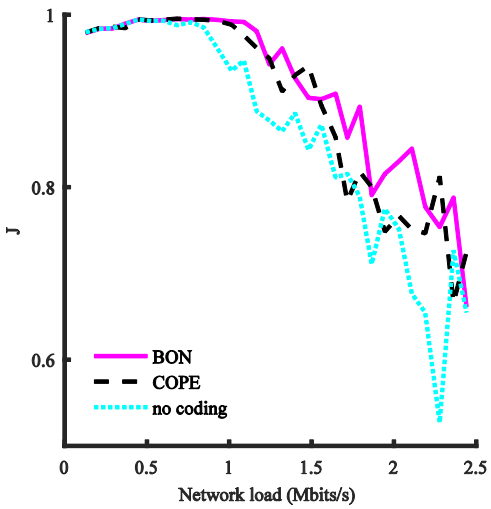


Figure 13: Jain's index (J) for the distribution with 20 traffic flows

We carried out a detailed analysis of the individual traffic flows with respect to how successful the network was in transferring individual flows from the source to the destination. Other parameters were set according to Table 1.

First, let us consider Jain's index, which is primarily used for identifying underutilized channels. In our case we are establishing whether users or applications were receiving a fair share of the system resources. In Fig. 13 Jain's index for 20 flows of traffic distribution is presented. Jain's index indicates that the network-coding procedures treat the traffic at least as well as or better than when there is no coding; however, it does not reveal how the coding gains were distributed among the flows.

In addition to Jain's index, in the following we use the gain of the particular traffic flow as the performance metric, so we refer to it as the flow gain. The flow gain results for a traffic distribution consisting of 20 traffic flows are shown in Fig. 14. The flow gain was collected at an overall network load equal to 0.768 Mbit/s, which is the maximum load at which the delay for the no-coding scenario was lower than 2 s. At such a network-load intensity, the goodput for the no-coding scenario was already affected by full queues on the nodes, i.e. the network was congested and several nodes were dropping packets due to the full queues. For the same traffic load, the COPE and BON network-coding procedures did not have a negative effect on any of the flows in terms of goodput, as in this case the highest observed flow gain was above 30%.

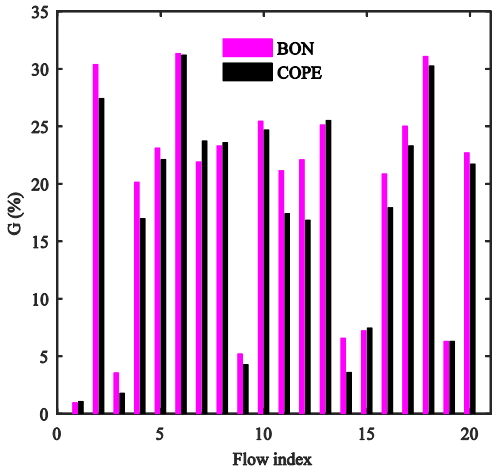


Figure 14: Gain (G) per individual flows for the traffic distribution with 20 traffic flows

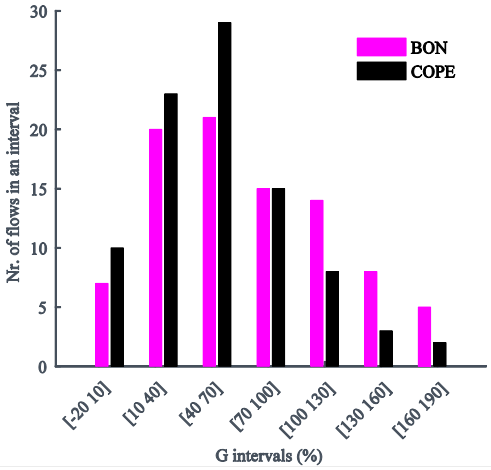


Figure 15: Histogram for gain (G) per individual flow for the traffic distribution with 90 traffic flows

In Fig. 15 we present the results for the case with an even higher number of traffic flows (i.e., 90). The results are presented as histograms for the overall network load equal to 1.24 Mbit/s, which was the maximum load at which the delay for the COPE was lower than 3 s. Again, the network was in a state where packets on several nodes were dropped due to the full output queues. On the x-axis we have the flow gain while on the y-axis is the number of flows within a given flow-gain interval. For example, there were 20 flows for which the BON achieved a flow gain between 10 % and 40 %. Although the goal of the network-coding procedures is to successfully transfer many flows with a high flow gain to their destinations, it is desirable that none of the flows are degraded. Both coding procedures were able to improve the goodput of almost all the flows and there were only a few flows where the gains were low. A detailed analysis revealed that there were 2 out of 90 flows with the BON having a negative coding gain, where the lowest value was -1.15 %. With the COPE there were four flows with a negative gain, with the lowest value being -16.7 %. Although we present only a few representative results, the same conclusions can also be drawn for different traffic distributions and higher traffic loads.

6 Conclusions

Two opportunistic network-coding algorithms the COPE and BON were used for evaluation purposes. Both procedures are similar in several aspects, but they differ in their core, i.e., the coding decision algorithm, which in case of the COPE using exact distributed information about coding opportunities, whereas in the case of the BON it is based only on the local information

available in a node. Both procedures show similar tendencies, thus observations and conclusions gained can be generalised to other similar types of algorithms. The performance of both coding algorithms could be further improved for specific scenarios by optimising the particular parameters or using additional support mechanisms. However, in this investigation we focused on the impact of the network and traffic characteristics on the efficiency of the network-coding procedures, and thus the network-coding parameters remained fixed throughout the evaluation.

In the following we summarise the observed and discussed design principles for wireless networks using network coding and their applications. Table 2 summarises the winning parameters and protocol with an included reference to the section where they are analysed.

Table 2: Summary of winning parameters for analysed scenarios

	Parameter	**Values**	**Win. value**	**Win. protocol**
Topology design (4.1)	Transmission range	280 m - 420 m	280 m	BON
Traffic distribution impact (4.2)	Traffic-symm. ratio	0% - 100%	$> 60\%$	BON
Traffic distribution impact (4.2)	Number of traffic flows	10 - 90, '*all2all*'	10	BON
Packet length Evaluation (4.3)	Packet length	2 kb - 10 kb, '*var*'	2 kb	BON
Analysis of queue length (4.4)	Queue length	10 p - 200 p	50 p	BON
QoS Analysis - ETE delay (5.1)	Number of traffic flows	2, 20	20	BON $\approx$ COPE
QoS Analysis - Jitter (5.1)	Number of traffic flows	2, 20	2	BON $\approx$ COPE
Coding Fairness Jain's index (5.2)	Number of traffic flows	20	20	BON $\approx$ COPE

Traffic distribution, and in particular, traffic symmetry have a significant impact on the performance of opportunistic network coding. In the case where only one-direction flows (completely asymmetric traffic) are considered, opportunistic network coding will not improve the overall network performance. By increasing the symmetry factor, the coding gains will increase, reaching the maximum benefits for a completely symmetric traffic load. Furthermore, the highest gains are obtained when there is a relatively small number of flows in the network. By increasing the number of flows the network-coding gains decrease.

The average length of the packets in the network is also an important parameter for the selection of the network-coding procedure. The overhead induced by the network-coding procedure has a larger influence on the network with smaller packets. When packets of the same length are encoded the gains obtained are higher than when the packets vary in length. In addition, the enlarging of the output queues improves the gains, but only to a certain level, which also depends on the number of flows in the network. Hence, the length of the output queue should primarily be adjusted to handle the expected peak network loads.

When planning the network topology, additional transmission power and thus an increased nodes-transmission range in network-coding-enabled networks will not necessarily improve the networks' performance. The results show that with the increased nodes' transmission range the benefits of opportunistic network coding fade, requiring cautious planning of the transmission power.

The detailed delay and jitter analysis of individual flows revealed no negative effects on the QoS due to the use of network-coding procedures. Although the jitter might have slightly increased (almost negligibly in absolute terms) on rare occasions with low loads, the network performance in terms of delay and jitter with network coding in a heavily loaded network significantly outperformed the no-coding scenario where no network coding was used. It is worth noting that the jitter and delay were analysed on a flow that travelled 14 hops to its final destination and that there were only a few occasions where packets were not encoded at least one time along the path, and still we did not detect any degradation in the QoS.

Finally, the coding-fairness analysis showed that both network-coding procedures considered in this study improved or at least maintained the throughput of all the traffic flows in the network,

i.e., they were not improving the overall performance at the expense of some less-favoured traffic flows.

Funding

This work has been funded by the Slovenian Research Agency through grants J2-4197 and P2-0016.

Bibliography

[1] Ahlswede, R.; Cai, N.; Li, S.-y. R.; Yeung, R. W. (2000). Network Information Flow, *IEEE Transactions on Information Theory*, 46, 1204-1216, 2000.

[2] Alic, K.; Pertovt, E.; Svigelj, A. (2012). Network coding simulation model in OPNET modeler, *In procc. of OPNETWORK 2012*, Washington, USA, 2012.

[3] Alic, K.; Pertovt, E.; Svigelj, A. (2015). Bearing-Opportunistic Network Coding, *International Journal of Computers, Communications & Control,* 10, 2015.

[4] Alic, K.; Svigelj, A. (2016). A One-Hop Opportunistic Network Coding Algorithm for Wireless Mesh Networks, *Wireless Networks, https://doi.org/10.1007/s11276-016-1384-y, 2016.*

[5] Alic, K.; Svigelj, A. (2017). Self-adaptive practical opportunistic network-coding procedure for static wireless mesh networks, *Ad Hoc & Sensor Wireless Networks,* 36, 87-105, 2017.

[6] Chieochan, S.; Hossain E. (2012). Network coding for unicast in a WiFi hotspot: Promises, challenges, and testbed implementation, *Computer Networks*, 56, 2012.

[7] Coppi, N. D.; Ning, J.; Papageorgiou, G.; Zorzi, M.; Krishnamurthy, S. V.; Porta, T. L. (2012). Network Coding Aware Queue Management in Multi-Rate Wireless Networks, *in Procc. of the 21st International Conference on Computer Communications and Networks (ICCCN)*, Munich, Germany, 2012.

[8] Hundeboll, M.; Ledet-Pedersen, J.; Heide, J; Pedersen, M. V.; Rein, S. A.; Fitzek, F. H. P. (2012). CATWOMAN: Implementation and Performance Evaluation of IEEE 802.11 Based Multi-Hop Networks Using Network Coding, *Procc. of the IEEE Vehicular Technology Conference (VTC Fall)*, Quebec City, Canada, 2012.

[9] Jain, R.; Chiu, D.-M.; Hawe W. (1984). *A Quantitative Measure Of Fairness And Discrimination For Resource Allocation In Shared Computer Systems.* Available: http://www.cse.wustl.edu/~jain/papers/fairness.htm

[10] Katti, S.; Rahul, H.; Hu, W.; Katabi, D.; Médard, M.; Crowcroft, J. (2008). XORs in the Air: Practical Wireless Network Coding, *IEEE/ACM Transactions on networking*, 16, 2008.

[11] Le, J.; Lui, J. C. S.; Chiu, D.-M. (2010); DCAR: Distributed Coding-Aware Routing in Wireless Networks, *IEEE Transactions on Mobile Computing*, 9, 596-608, 2010.

[12] Matsuda, T.; Noguchi, T.; Takine, T. (2011). Survey of Network Coding and Its Applications, *IEICE Transactions on Communications*, E94-B, 698-717, 2011.

[13] Omiwade, S.; Zheng, R.; Hua, C. (2008). Butteries in the Mesh: Lightweight Localized Wireless Network Coding, *Procc. of Fourth Workshop on Network Coding, Theory and Applications (NetCod)*, Hong Kong, 2008.

[14] Pertovt, E.; Alic, K.; Svigelj, A.; Mohorčič, M. (2018). CANCAR - Congestion-Avoidance Network Coding-Aware Routing for Wireless Mesh Networks, *KSII Transactions on Internet and Information Systems*, 12, 9, 2018.

[15] Seferoglu, H.; Markopoulou, A. (2009). Video-aware opportunistic network coding over wireless networks, *IEEE Journal on Selected Areas in Communications,* 27, 713-728, 2009.

[16] Seferoglu, H.; Markopoulou, A., Medard, M. (2011). NCAPQ: Network Coding-Aware Priority Queueing for UDP Flows over COPE, *In procc. of the Int. Symp. on Network Coding (NetCod)*, Beijing, China, 2011.

[17] Seferoglu, H.; Markopoulou, A. (20013). Network Coding-Aware Queue Management for TCP Flows Over Coded Wireless Networks, *IEEE/ACM Transactions on Networking,* 2013.

[18] Svigelj, A.; Mohorcic, M.; Kandus, G.; Kos, A.; Pustisek, M.; Bester, J. (2004). Routing in ISL Networks Considering Empirical IP Traffic, *IEEE Journal on Selected Areas in Communications* , 22, 2004.

[19] Wang, S.-Y.; Lin, C.-C.; Chang, Y.-C. (2012). A rule-based inter-session network coding scheme over IEEE 802.16(d) mesh CDS-mode networks, *Computer Networks*, 56, 661-685, 2012.

[20] Yan, Y.; Zhang, B.; Zheng, J.; Ma, J.(2010). CORE: a coding-aware opportunistic routing mechanism for wireless mesh networks, *IEEE Wireless Communications*, 17, 96-103, 2010.

[21] You, L.; Ding, L.; Wu, P.; Pan, Z.; Hu, H.; Song, M. (2011). Cross-layer optimization of wireless multihop networks with one-hop two-way network coding, *Computer Networks,* 55, 1747-1769, 2011.

[22] Zeng, D.;Guo, S.; Leung, V.; Hu, J. (2011). The Exploration of Network Coding in IEEE 802.15.4 Networks, *Int. Journal of Digital Multimedia Broadcasting,* vol. 2011, 2011.

[23] Zhang, J.; Chen, Y. P.; Marsic, I. (2010). MAC-layer proactive mixing for network coding in multi-hop wireless networks, *Computer Networks,* 54, 196-207, 2010.

[24] Zhao, F; Medard, M. (2010). On analyzing and improving COPE performance, *Proc. of Information Theory and Applications Workshop*, Italy, 2010.

[25] Zhao, F; Medard, M.; Hundeboll, M.; Ledet-Pedersen, J.; Rein, S. A.; Fitzek, F.H.P. (2012). Comparison of Analytical and Measured Performance Results on Network Coding in IEEE 802.11 Ad-Hoc Networks, *Int. Symp. on Network Coding (NetCod), 2012*, USA, 2012.

INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL
ISSN 1841-9836, e-ISSN 1841-9844, 14(3), 311-328, June 2019.

Research on Key Technology of Web Hierarchical Topic Detection and Evolution Based on Behaviour Tracking Analysis

Mo Chen*
Business College of Beijing Union University
A3, Yanjingdongli, Chaoyang District, Beijing, 100025, P.R. China
*Corresponding author: mo.chen@buu.edu.cn

Abstract: In the development background of today's big data era, the research direction of Web hierarchical topic detection and evolution characterized by the semi-structured or unstructured data has caught wide attention for academicians. This paper proposes an idea of Web hierarchical topic detection and evolution based on behaviour tracking analysis taking the network big data as the research object, and expounds main implementation methods, which include the instance analysis of the usage mode, the instance analysis of the seed, the set analysis of similar instance supporting the topics, the set analysis of similar instance supporting the events, the evolution analysis of the event, and expounds the algorithm of Web hierarchical topic detection and evolution based on behaviour tracking analysis. The process of experimental analysis is organized as follows, first of all, the experiment analyses the quality of topic detection, the accuracy rate with the number of instance concerned and the seed threshold variation trend, the accuracy rate with the number of instance concerned and the probability threshold variation trend, secondly, the experiment analyses the quality of topic evolution, the accuracy rate with the variation trend of parameter adjustment, the accuracy rate with the number of instance concerned and the similar threshold variation trend, finally, the experiment analyses the time consuming to solve main research problem under different method, the qualitative result of topic detection and evolution under different data set. The results of experimental analysis show the idea is feasible, verifiable and superior, which plays a major role in reconfiguring Web hierarchical topic corpus and providing an intelligent big data warehouse for the network information evolution application.
Keywords: Web hierarchical topic, topic detection, event evolution, behaviour tracking analysis.

1 Introduction

In the development background of Web text mining technology and big data era, so far, the field of intelligent technology has also developed into a more challenging stage [7, 13, 28], the network has become one of services, which can transmit most popular information for users. According to deep survey, the number of network data has gone through EB level in different domain [14, 18, 19, 27]. Academicians should cogitate how to analyse intricate big data, nevertheless, it is an important and key application direction for researching Web hierarchical topic detection and evolution based on behaviour tracking analysis.

In the network big data, the number of Internet news is showing explosive growth as a kind of flow resource with on-going events, which has shown 5V features of volume, variety, value, velocity and veracity [3, 23, 26]. Based on above characteristics, the Internet news should reflect high currency and reliability [5], on this basis, the topic of Internet news should be quickly detected, and its evolution path should be tracked in real time. However, how to research Web hierarchical topic detection and evolution based on behaviour tracking analysis, it has become an urgent problem to build a Web hierarchical topic corpus and provide a real-time big data source for the network information evolution application.

Through researching the literature related to topic detection and evolution technology, this paper proposes an idea of analysing the process for Web hierarchical topic detection and evolution, and expounds main implementation methods and algorithms following with interest Web hierarchical topic detection and evolution from behaviour tracking analysis. This process does important contribution for researching a method of analysing the detection and evolution for Web hierarchical topic, the results of experimental analysis show that the implement of this idea is feasible, verifiable and superior.

2 Related works

In recent years, some scholars have done certain research about the technology of Web topic detection and evolution. A statistical model is proposed [2], in this model, it can combine context with related topics by jointly modelling the topic word with the hash tag and the time stamp, in order to detect and track interpretable topics over time along with their distribution of the hash tag, in this technical context, the experiment demonstrates that this model effectively reveals the process of topic detection and evolution by using the real dataset, this model is different from the traditional topic mining model, it shows serious improvement due to this fact that the distribution of the metadata containing in user content generated can be analysed, so the whole research result does main contribution in the area of topic detection and evolution in the context of the statistical analysis. A topic detection and evolution method is proposed by analysing the semantic word shift, the topic trend, and the evolving dynamic using the data set [4], in this method, it can merge and split local topics in different time periods, in order to track the process of knowledge transfer among topics, in this technical context, the experimental results show that the process of topic detection and evolution usually follows pattern from adjusting status to mature status, and sometimes with readjusting status, this method is different from the statistical analysis, it shows serious improvement due to this fact that the word migration via topic channels has been defined, and three migration types of non-migration, dual-migration, and multi-migration are better to understand topic detection and evolution, so the whole research result does main contribution in the area of topic detection and evolution in the application direction of information retrieval. A topic detection and evolution method is proposed [30], which is called the citation-content-latent Dirichlet allocation method, in this method, it can account for the document citation relation and the content of document itself via a probabilistic generative model, this model can deal with the citation and text information, and its parameters are estimated by a collapsed Gibbs sampling algorithm, in addition, a topic detection and evolution algorithm is designed, which can run in two steps of the topic segmentation and the topic dependency relation calculation, this model and algorithm have been tested by using the online dataset, in this technical context, the experimental results demonstrate that the implementation of the model and algorithm can more effectively detect important topics and reflect topic evolution process comparing with the topic tracking in the knowledge transfer context, so the whole research result does main contribution in the area of topic detection and evolution for designing the model and algorithm. A topic detection and evolution framework is proposed based on the probabilistic topic model [31], in this framework, firstly, the notations, the terminology, and the basic topic mining model is introduced, secondly, three technologies of the topic detection and evolution are applied, which are the discrete time topic detection and evolution, the continuous time topic detection and evolution, and the online topic detection and evolution, thirdly, the application of this framework is discussed, in this technical context, the comparative experiments are completed for different technologies of the topic mining, this framework shows serious improvement than single probabilistic model and does main contribution in the area of the topic detection and evolution performance evaluation. A topic detection and evolution method is proposed based on the analysis of the content similarity

or dissimilarity using the textual material [12], in this method, the graph-theoretical technology is applied, in order to deal with the network relationship among the content-similar topics, in this technical context, the explanatory experiment more effectively illustrates usefulness of the approach using the online news articles in different situations, so the whole research result does main contribution in the area of topic detection and evolution in the context of the network analysis.

Based on above analysis of the literature about the technology of Web topic detection and evolution, through comparing the difference among technologies and analysing the main contributions in the research area, if want to research the process of Web topic detection and evolution, in addition to using Web mining technology based on the structure and content, but also using Web mining technology based on the behaviour tracking. Therefore, in recent years, some scholars have also done certain research about the technology of Web behaviour tracking analysis.

A collaborative tagging model is proposed [24], in this model, the technology of the usage behavior tracking analysis is applied to the information extraction direction for web user query in the ontology environment due to different structure data, which is based on the idea of the block acquiring page segmentation, in order to retrieve the tag-based information, in this technical context, the comparative experiments are completed regarding the average precision rate, the time cost, and the storage space rate with existing information retrieval model, it shows that the application of this model can assure research more effectiveness, so the whole research result does main contribution in the area of the behavior tracking analysis. A tracking method of the usage behavior is proposed [17], in this method, a relational graph is established by mining the temporal and causal information among aggregated HTTP request, and an algorithm is designed and implemented for primary request identification, which is a critical task of web usage mining, in order to demonstrate higher value and effectiveness, in this technical context, the experimental result shows that it is a more useful method of analysing a large-scale dataset for the real-world Web access log, so the whole research result does main contribution in the direction of mining value for information available. A web usage mining method is proposed [1], this method can be applied to solve the problem of developing more accurate and efficient recommendation systems, the traditional data protection mechanism focuses on the access control and the secure transmission, which provide only security against malicious the third parties, but not the service provider, so this method can mine efficiently and intelligently data, in order to guide and track the users' usage behavior, and does main contribution in the application direction of the modern E-business. A web usage mining method is proposed [11], in this method, the state-of-the-art session identification technology is used in terms of limitation, feature, and methodology, which provide a structured overview of the research development, in this technical context, the comparative experiments critically review existing session identification technology, the whole research result does main contribution in highlighting the limitation and related challenge, identifying the area where further improvement is required, in order to complement the performance of existing technology. A web usage mining method is proposed [21], in this method, an algorithm is also designed for the data cleaning and filtering using the web log dataset, this algorithm mainly complete the process of preprocessing and clustering the usage behavior, which consists of the use cases for the data cleaning and filtering, the user and session identification, in this technical context, the experiments are carried out to obtain the aggregate clustering results, through this process, two datasets of web usage log are collected and processed, so the whole research result does main contribution in the area of web usage behavior analysis.

Based on above analysis of the literature about the technology of Web behavior tracking analysis, through comparing the difference among technologies and analysing the main contributions in the research area, the scholars have studied two research directions including the

technology of Web topic detection and evolution, but do not fully take into account the hierarchical series induced by the usage behavior, if do not fully consider this point, then ignore the process to track the usage behavior from the angle of topic detection and evolution. So this paper mainly takes web usage behaviour record of news big data as the research source, utilizes the method of analysing web usage behaviour tracking to perfect news big data corpus for topic detection and evolution research, and proposes an idea of analysing the process for Web hierarchical topic detection and evolution to solve difficult problems existing in current research status.

3 Problem definition and notation

Under the background of Web big data development, users can retrieve and browse Web news from different dimension, granularity and frequency, which has been analysed and evaluated [15,20,22]. In this process, the time sequence trajectory of users' behavior can be recorded. These data not only record the characteristic of Web news that users use, but also contain the topics reflected in Web news, and the events that are generated based on these topics. Therefore, the knowledge hidden in Web news big data can be mined, the topics that users are concerned about can be detected, a series of events under the topics can be tracked, and the evolution process of events can be combed out based on the process of analysing Web news usage characteristic.

Based on the analysis of Web news structure, contents and semantic feature, every Web news that users concern can be regarded as an instance node in every authoritative Web news network from the perspective of global usage. The social events supported by set of some related nodes can be considered as a topic, and a series of events will be created under each topic. In this way, when users are concerned about a series of topics reflected in a social event, they can not only browse many Web news instances supporting topics, but also browse a series of events that are generated by this topic. When users are concerned about an event, it can also browse more than one Web news instance content that supports this event. From the perspective of local usage, when users retrieve Web news instances, besides inputting the keywords related to social events reported by Web news, they can also input the keywords with five tuple semantic description. Therefore, during the analysis process of Web news topic detection and evolution, the social events can be mined utilizing structure, contents and semantic feature, a series of events caused by the topics can be mined focusing on mining Web news instances supporting the events, the logical hierarchical relationship can be mined between mining objects, which will construct a multi-level structural corpus, it can represent visual topics and events of Web news with a high degree of quality.

Based on the analysis of Web news utility feature, users can retrieve Web news in the search process with a high degree of currency, which is called the time accuracy reporting Web news core events. Users can retrieve Web news with a high degree of truthfulness, which is called the releasing source reliability of Web news. Users can retrieve Web news of high currency with a high degree of truthfulness, therefore, during the analysis process of topic detection and evolution, it should weigh the factors of Web news currency and authenticity, and provide Web news utility instances supporting the topics and events from the perspective of users' utility characteristic for Web news.

Based on the behaviour tracking analysis of web usage characteristic for user, $S-U$ can link the search keywords and the URL instances synchronously, in the $S-U$ relation, S represents the set of the keywords, U represents the set of the URL instances. As shown in the formula 1, $fq(s,u)$ can indicate the clicking frequency of instances, $fq_i(u)$ can indicate the clicking frequency of homologous URL, $fq(u)$ can indicate the clicking frequency of instances in a certain period. As shown in the formula 2, $rt_i(u)$ can indicate the clicking rate of homologous URL.

$$fq(u) = \sum_{i=1}^{n} fq_i(u) \tag{1}$$

$$rt_i(u) = \frac{fq_i(u)}{fq(u)} \tag{2}$$

$NewsSet$ can be defined using the $\{ns_1, ..., ns_{i-1}, ns_i, ns_{i+1}, ..., ns_k\}$, the range of array is from one to k. $ns_i.url$ defines the url instance address, $ns_i.title$ defines the instance title, $ns_i.pubtime$ defines the instance releasing time, $ns_i.pubsource$ defines the instance releasing source, $ns_i.content$ defines the instance contents, $ns_i.keyword$ defines the instance keywords. $UserBehavior$ can be defined using the $\{ub_1, ..., ub_{i-1}, ub_i, ub_{i+1}, ..., ub_n\}$, the range of array is from one to n. $ub_i.username$ defines the user name, $ub_i.searchword$ defines the keywords, $ub_i.url$ defines the URL clicked, $ub_i.systemtime$ defines the system time.

Based on above definition and notation, the issue should be solved to mine topics contained in the instances, mine instances set supporting related topics. From research angle of mining Web news instance set supporting topics concerned by users, the issue should be solved to mine events that are happening under these topics, analyse the instance set supporting these events, so as to reflect the evolution process of Web news topics continuously. This result can be denoted using the $TopicURL$, it can be represented using the $\{tu_1, ..., tu_{i-1}, tu_i, tu_{i+1}, ..., tu_m\}$, the range of array is from one to m, tu_i can be represented using the $< Topic, Topicurl, Event, Eventurl >$, $tu_i.Topic$ can express the topic description detected, $tu_i.Topicurl$ can indicate the seed instances URL set supporting related topics, $tu_i.Event$ can express the description of events under topics mined, $tu_i.Eventurl$ can indicate Web news instance of events supporting topics detected.

4 The analysis of Web hierarchical topic detection and evolution

In view of the problem definition and notation, this paper proposes an idea of analysing the process for Web hierarchical topic detection and evolution shown in figure 1. This framework is used for completing the analytical process for Web hierarchical topic detection and evolution, which include the instance analysis of the usage mode, the instance analysis of the seed, the set analysis of similar instance supporting the topics, the set analysis of similar instance supporting the events, and the evolution analysis of the event. The algorithms are designed for completing the functions and methods of this framework, which include the algorithm of analysing the process for Web hierarchical topic detection, the algorithm of analysing the process for Web hierarchical topic evolution.

4.1 The algorithm of analysing the process for Web hierarchical topic detection

The algorithm of analysing the process for Web hierarchical topic detection is implemented by designing two methods of the usage mode analysis and the topic series construction, the inputting content of this algorithm is the result of semantic five tuple description analysis and utility evaluation for Web news instances, and the user usage behaviour record set, the outputting content of this algorithm is the similar and sequential set of Web news instances that can support corresponding topics.

According to web usage behaviour record, the explosive and attention mode of Web news instances are analysed, in order to infer the click mode of Web news instances, the degree distribution and similarity mode of Web news instances are also analysed, in order to infer the retrieval mode of Web news instances. In accordance with these results of analysing web usage

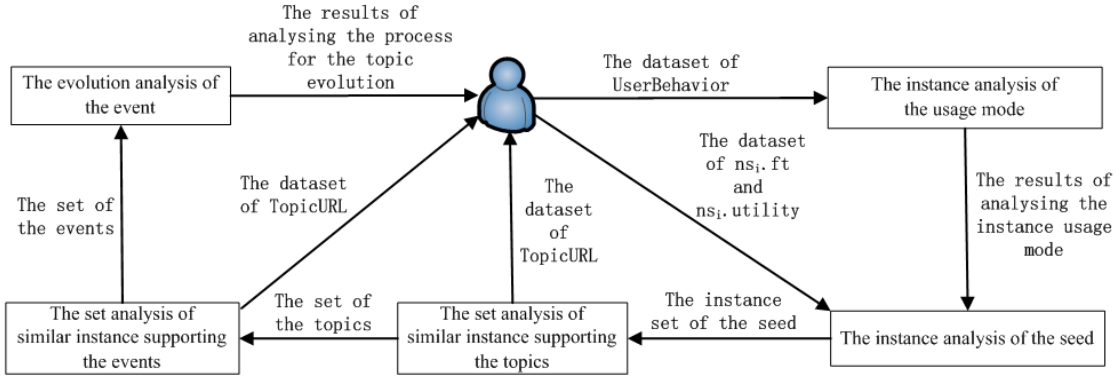


Figure 1: The framework of analysing the process for Web hierarchical topic detection and evolution

mode, the seed set of Web news instances can be mined, the similar set of Web news seed instances can also be mined. Referring to the utility feature that have been analysed [6], the semantic five tuple can describe the topics under the time series.

For the process of analysing the explosion mode, the execution process can be viewed a sensor for the social event. The process of analysing the mode quotes the entropy characteristic, and the analytical result is a speculation about the sharpness of the click rate change. For the process of analysing the attention mode, the execution process makes up for the problem existing in the burst mode, that is, when the research instances are followed by web users, the measurement standard will be an absolute phenomenon. Based on the analysis of the usage mode, the click mode of Web news instances can be inferred shown in the formula 3.

$$ClickMode(u) = (1 - (-\sum_{i=1}^{n} rt_i(u) \times log_n rt_i(u))) \times \\ \times \frac{log(fq(u)) - Min_{u_i \in U}(log(fq(u_i)))}{Max_{u_i \in U}(log(fq(u_i))) - Min_{u_i \in U}(log(fq(u_i)))} \tag{3}$$

In the formula 3, n indicates the number of the granular unit that the instances are followed. For the sudden event, it can be set in day. For the normal occurrence event, it can be set in week or month. If the fluctuation is not large for the click rate of Web news instances, then the attention mode is smaller. If Web news instances have obvious fluctuation, then the attention mode will be large. According to the click process of Web news instances, it has the power law distribution characteristic. Therefore, the click frequency of Web news instances has carried on logarithm transformation.

For the process of analysing the degree distribution mode, the degree of Web news instances presents the power law distribution, so its logarithmic transformation can be executed. For the process of analysing the similarity mode, it not only makes up for the problem existing in the degree distribution mode, which ignores the degree origin of Web news instances through retrieval keyword, but also solves the problem of the sparse record in the click behaviour of web user. Based on the analysis of the usage mode, the retrieval mode of Web news instances can be speculated shown in the formula 4.

$$SearchMode(u) = \frac{2}{n(n+1)} \times \frac{log(d(u)) - Min_{u_i \in U}(log(d(u_i)))}{Max_{u_i \in U}(log(d(u_i))) - Min_{u_i \in U}(log(d(u_i)))} \times \times \sum_{i \leq j}^{n} \frac{\sum_{k \in dataitem}^{\infty}(s_{ik}(u)s_{jk}(u))}{\sqrt{\sum_{k \in dataitem}^{\infty}(s_{ik}(u))^2}\sqrt{\sum_{k \in dataitem}^{\infty}(s_{jk}(u))^2}} \quad (4)$$

Based on the formula 3 and 4, the set of Web news instances can be mined by using the formula 5, $SeedURL(u)$ should be larger than or equal to the seed threshold. As shown in the equation 6, $< t_i, p_i, o_i, ce_i, re_i > (ns_i.uc)$ represents the semantic five tuple description for the seed instance. For the utility evaluation result, the sort sequence of the utility can be executed. In the subsequent experiments, the optimal value of the seed threshold will be analysed.

$$SeedURL(u) = ClickMode(u) \times SearchMode(u) \quad (5)$$

$$TimeSeries(ns) = (< t_1, p_1, o_1, ce_1, re_1 > (ns_1.uc), ..., < t_i, p_i, o_i, ce_i, re_i > (ns_i.uc), ..., < t_n, p_n, o_n, ce_n, re_n > (ns_n.uc)) \quad (6)$$

Algorithm 1 Method 1 Analysing the Usage Mode

```
Input: UserBehavior, Threshold;
Output: TopicURL;
LET UserRecord ← UserBehavior;
LET GroupUserRecord ← GroupByURL(u), TopicURL ← φ;
For each gur[i](0 ≤ i ≤ gur.size() − 1) Do
    SeedURL ← Calculate clickmode and searchmode;
    If SeedURL ≥ Threshold Then
        tu.add(SeedURL);
    End If
End For
```

For the construction of the sequence topic, the execution process uses the probability for its first transfer, in order to judge whether it is similar to the seed instance supporting topics by using Web news instances and taking Web news seed instance as the research center. If su indicates the seed instance, then the variable t_u expresses that whether the instance can support the topic of su, the variable t_s expresses that whether the search keyword can support the topic of su. If the seed instance is able to support the topic of su, then $t_u = 1$, conversely, $t_u = 0$, if the search keyword is able to support the topic of su, then $t_s = 1$, conversely, $t_s = 0$. In initial status, t_{su} is one, $P(t_{su} = 1)$ is one, the probability is zero. As shown in the formula 7 and 8, $P(t_s = 1)$ is able to calculate su probability, $P(t_u = 1)$ is also able to be recalculated, when $P(t_u = 1)$ is larger than or equal to the probability threshold, the Web news instances are able to be found, so as to mine the similar instance set. In the subsequent experiments, the optimal value of the probability threshold will be analysed.

$$P(t_s = 1) = \sum_{u:(s,u) \in E}^{\infty} \frac{fq(s,u)}{\sum_{(s,u_i) \in E}^{\infty} fq(s,u_i)} \times P(t_u = 1) \quad (7)$$

$$P(t_u = 1) = \sum_{s:(s,u)\in E}^{\infty} \frac{fq(s,u)}{\sum_{(s_i,u)\in E}^{\infty} fq(s_i,u)} \times P(t_s = 1) \tag{8}$$

Algorithm 2 Method 2 Constructing the Topic Series

```
TopicURL, UserBehavior, Threshold;
TopicURL;
For each tu[i](0 ≤ i ≤ tu.size() − 1) Do
  swset1 ← ExistSet(tu[i].getSet("Topicurl"), ub);
  While swset1 is not null Do
    While each swset1 is exist Do
      p(t_s) ← CalculateResult(swset1.getElement(j).position, tu[i].getSet("Topicurl"), ub);
      If p(t_s) ≥ Threshold Then
        swset2.addSet(swset1.getElement(j));
      End If
      ub ← (swset1.getElement(j).position, p(t_s));
    End While
    While each swset2 is exist Do
      wnuset1 ← ExistSet(swset2.getElement(j).position, ub);
      If(wnuset1 ← EqualSet(wnuset1, wnuset2)) is not null Then
        While each wnuset1 is exist Do
          p(t_u) ← CalculateResult(wnuset1.getElement(k).position,
                                   swset2.getElement(j).position, ub);
          If p(t_u) ≥ Threshold Then
            wnuset2.addSet(wnu1.getElement(k));
          End If
          ub ← (wnuset1.getElement(k).position, p(t_u));
        End While
      End If
    End While
    swset1 ← ExistSet(wnuset2, swset2, ub);
  End While
  tu[i] ← wnuset2;
  Describe Topic tu[i];
End For
```

4.2 The algorithm of analysing the process for Web hierarchical topic evolution

The algorithm of analysing the process for Web hierarchical topic evolution is implemented by designing two methods of the event series construction and the event evolution analysis, the inputting content of this algorithm is the result of analysing semantic five tuple for Web news instances, the user usage behavior record set and Web news topic set, the outputting content of this algorithm is the topic set that has been excavated under the events and the evolution result of analysing the events belonging to the topics.

According to the set of Web news instances that can support the topics mined, the user usage behavior record is used to analyse the time sequence of Web news instances that are followed, the similarity degree of the core events reported among Web news instances is calculated by

using the result of analysing semantic five tuple description, and the evolution state of the events supported by the similar Web news instances and the topics belonging to the events can also be analysed.

The calculation result of the time series similarity can show the topics mined among Web news instances, the user can pay much attention to Web news instances that describe the same events, which have occurred within a certain time period, and the time sequence concerned is similar. On a certain granularity, the vector of the time sequence can be expressed as shown in the formula 9 representing the attention rate of Web news instances. In this formula, $rt_i(u_j)$ represents the attention rate of the instance u_j in i component for a granularity, n indicates the number of the granular units that the instances are continuously followed. For the sudden events, the granularity can be set in days, for the normality events, the granularity can be set in weeks or months.

$$\begin{aligned} TimeSeries(tu_i) = (\{rt_{11}(tu_{11}), ..., rt_{1j}(tu_{1j}), ..., rt_{1m}(tu_{1m})\}, ..., \{rt_{i1}(tu_{i1}), \\ ..., rt_{ij}(tu_{ij}), ..., rt_{im}(tu_{im})\}, ..., \{rt_{n1}(tu_{n1}), ..., rt_{nj}(tu_{nj}), ..., rt_{nm}(tu_{nm})\}) \end{aligned} \tag{9}$$

According to the semantic five tuple description of Web news instances supporting the topics, the core events can be extracted, the similarity degree can be calculated among the core events reported by Web news instance. As shown in the formula 10, $FS(fs_i, fs_j)$ can be calculated with a threshold, which is greater than or equal to the similarity threshold. The semantic five tuple is used to describe the core events reported by the aggregated Web news instance set, the Web news instances that represent the node of each event are arranged in ascending order according to the occurrence time of the core events. If the core events occur in the same time, then Web news instances are arranged in descending order according to its utility characteristic. The Web news instances that are clustered together can be arranged in ascending order according to the occurrence time of the core events, if the core events occur in the same time, then Web news instances are arranged in descending order according to its utility characteristic. $< t_i, p_i, o_i, ce_i, re_i >$ $(< T_i, E_i > .uc)$ expresses the seed topic and the event description, $< t_{ij}, p_{ij}, o_{ij}, ce_{ij}, re_{ij} >$ $(< T_i.uc >, < E_{ij}.uc >)$ expresses the seed topic of the evolution event description. In the subsequent experiments, the optimal range of analysing the parameters and the similarity threshold value will be analysed.

$$\begin{aligned} FS(fs_i, fs_j) = \alpha \times \frac{\sum_{m=1}^{n}(ts_{im}, ts_{jm})}{\sqrt{\sum_{m=1}^{n}(ts_{im})^2}\sqrt{\sum_{m=1}^{n}(ts_{jm})^2}} + \\ +\beta \times \frac{\sum_{m=1}^{n}(ce_{im}, ce_{jm})}{\sqrt{\sum_{m=1}^{n}(ce_{im})^2}\sqrt{\sum_{m=1}^{n}(ce_{jm})^2}} \end{aligned} \tag{10}$$

$$\begin{aligned} TopicURL(T, E) = (\{< t_{11}, T_1, E_{11} >, ..., < t_{1j}, T_1, E_{1j} >, ..., < t_{1m}, T_1, E_{1m} >\}, ..., \\ \{< t_{i1}, T_i, E_{i1} >, ..., < t_{ij}, T_i, E_{ij} >, ..., < t_{im}, T_i, E_{im} >\}, ..., \{< t_{n1}, T_n, E_{n1} >, ..., \\ < t_{nj}, T_n, E_{nj} >, ..., < t_{nm}, T_n, E_{nm} >\}) \end{aligned} \tag{11}$$

5 The experimental analysis and result

In the process of completing the experiments based on designing the algorithms, the experimental environment of the software and hardware is used as follows. Java language is used for the

Algorithm 3 Method 3 Constructing the Event Series

```
Input: TopicURL, UserBehavior, NewsSet, Threshold, Parameters;
Output: TopicURL;
For each t[i](0 ≤ i ≤ t.size() − 1) Do
    Generate timeseriesvector;
    Fs ← Calculate similarity of timeseries and coreevent;
    If Fs ≥ Threshold Then
        Adjust event under topic t[i];
    End If
    Describe event under topic t[i];
End For
```

programming design to implement the algorithms, MyEclipse platform of the software research and development is used for the framework implementation, SQL Server of the database management system is used for web big data storage and process. The processor is Intel 2.40GHz, the memory is 32GB [**?**, 8, 9, 16, 29]. The experiments mainly use the standard data set for the social event of German A320 airliner crash, the data source is from massive Web news analysis corpus for the real data, the experimental analysis and result can verify feasibility and effectiveness of the research idea.

5.1 The qualitative analysis of the topic detection

As shown in the figure 2, the accuracy rate represents the quality of the topic detection by using three web usage behaviour mining processes. Firstly, the red column represents the accuracy rate of analysing the instance clicking mode, this quality is not high, although it has improvement, but the maximum is able to only arrive on about 0.64. Secondly, the blue column represents the accuracy rate of analysing the instance searching mode, this quality is not also high, although it has also improvement in several monitoring points, but the maximum is able to also only arrive on 0.63. Thirdly, the green column represents the accuracy rate of the algorithm designed in this paper, this quality is significantly improved, and the maximum is able to arrive on about 76

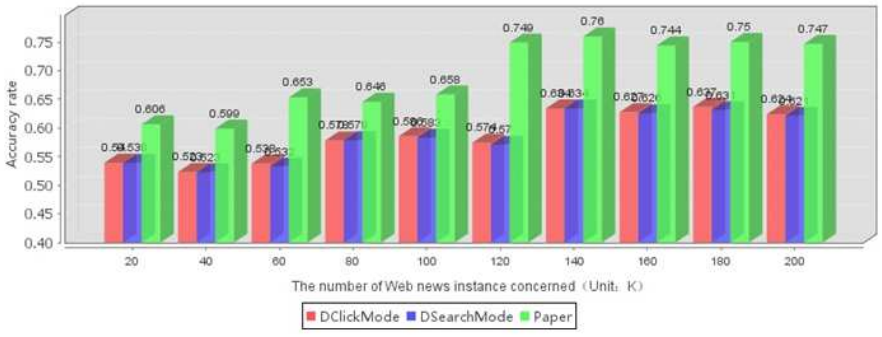


Figure 2: The qualitative analysis of the topic detection

5.2 The analysis of the accuracy rate with the number of the instances concerned and the seed threshold variation trend

As shown in the figure 3, the accuracy rate of the topic detection is indicated through the X and Y axis adjustment. The accuracy rate represents the quality of the topic detection, if the seed threshold is defined, then it is able to increase in a stable trend, because the number of the instances followed is less, the relationship among big data is simpler. If the number of the instances is increasing, then the relationship of the link exists, so the accuracy rate of the topic detection is increasing in a stable trend. If the number of the instances is defined, firstly, then the quality of the topic detection expresses an increasing trend, secondly, then it will decrease with the increasing threshold, because the threshold is less, the topics of inaccurate accuracy are able to be found. If the threshold can arrive at a stable range, then the topics of approximate accurate are able to be found. If the threshold can arrive at a value, then the accurate topics cannot be found. This experiment expresses when the number of the instances is one hundred and sixty, and the seed threshold is zero point seven five, the quality of the topic detection can get the highest about 0.78.

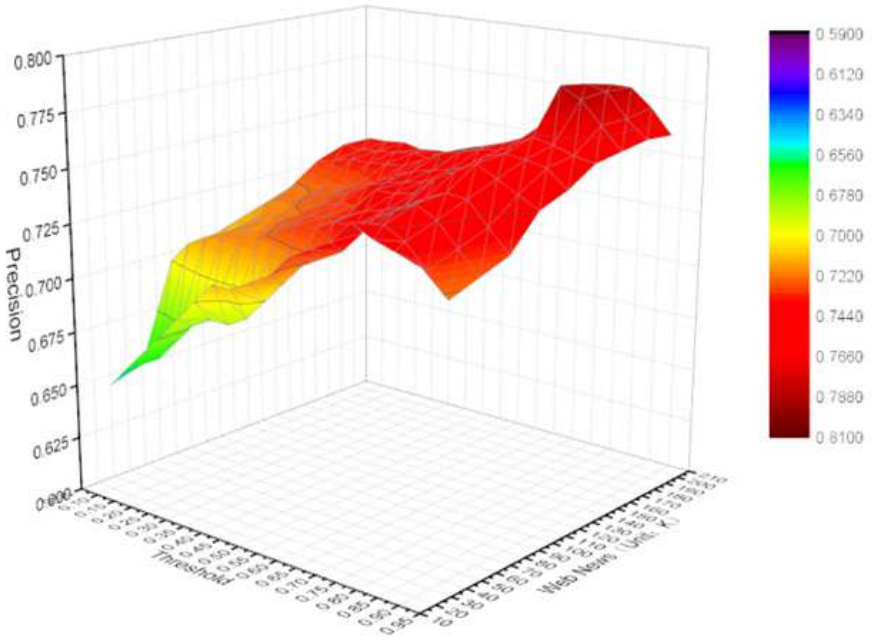


Figure 3: The analysis of the accuracy rate with the number of the instances concerned and the seed threshold variation trend

5.3 The analysis of the accuracy rate with the number of the instances concerned and the probability threshold variation trend

As shown in the figure 4, the accuracy rate represents the quality of the instances mined that support the topics by adjusting the probability threshold of the X axis and the number of the instances of the Y axis. The accuracy rate indicates the quality of the instances mined, if the threshold is defined, then the quality of the instances mined is able to increase in a stable trend, because the number of the instances followed is less, the relationship among big data is simpler. If the number of the instances concerned is increasing, then the relationship of the link exists, so the accuracy rate of the instances mined is increasing in a stable trend. If the number

of the instances is defined, firstly, then the quality of the instances mined expresses an increasing trend, secondly, then it will decrease with the increasing threshold, because the threshold is less, the instances of inaccurate accuracy can be mined. If the threshold can arrive at a stable range, then the instances of approximate accurate can be mined. If the threshold can arrive at a value, then the accurate instances cannot be mined. This experiment expresses when the number of the instances followed is one hundred and forty, and the probability threshold is zero point seven, the quality of the instances mined can get the highest about 0.76.

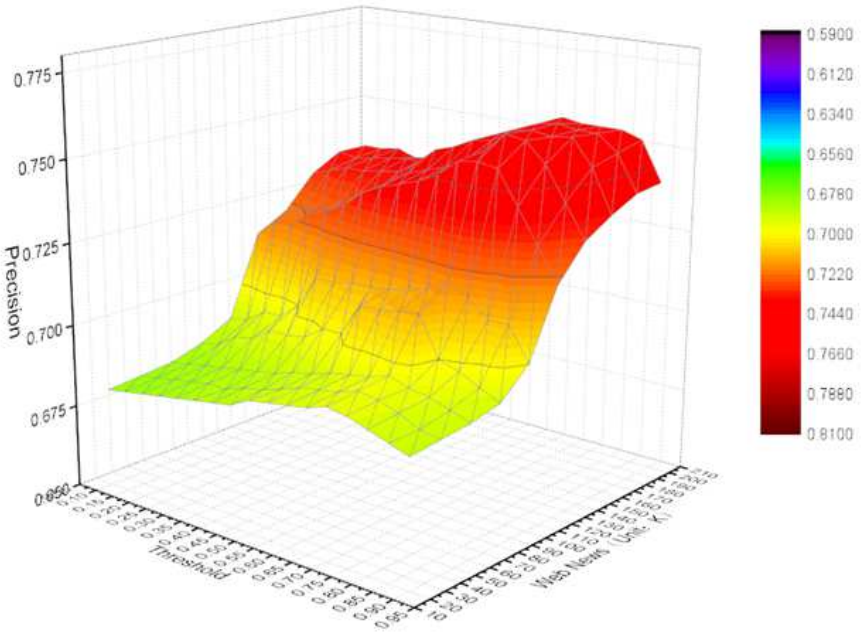


Figure 4: The analysis of the accuracy rate with the number of the instances concerned and the probability threshold variation trend

5.4 The qualitative analysis of the topic evolution

As shown in the figure 5, the accuracy rate represents the quality of the topic evolution by using three web usage behaviour mining processes. Firstly, the red solid line represents the accuracy rate of analysing the time series similarity, which shows that the accuracy rate is not high with the increase in the number of Web news instances concerned, although it has risen, but the highest can only arrive on about 0.64. Secondly, the blue solid line represents the accuracy rate of analysing the core event similarity, which shows that the accuracy rate is not also high with the increase in the number of Web news instances concerned comparing with analysing the time series similarity, and the accuracy rate has also a little decreasing slightly trend, the highest can only arrive on about 0.64. Thirdly, the green solid line represents the accuracy rate of the algorithm designed in this paper, which shows that the accuracy rate has greatly improved because of integrating the time series similarity based on the semantic evaluation and the similarity analysis for the core events. Although the accuracy rate is similar comparing with other two methods under the circumstance of less Web news instances concerned, but the accuracy rate has gradually widening the gap comparing with other two methods with the increase in the number of Web news instances concerned, the highest can arrive on about 0.75. So this experiment expresses that the quality of analysing the topic evolution is higher than other two methods by using the algorithm designed in this paper.

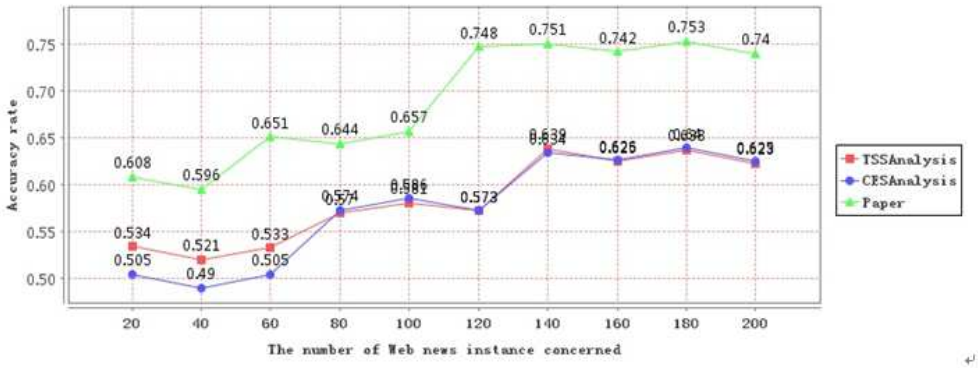


Figure 5: The qualitative analysis of the topic evolution

5.5 The accuracy rate of analysing the topic evolution with the variation trend for the parameter adjustment

As shown in the figure 6, the accuracy rate indicates the quality of analysing the topic evolution for Web news topics according to the parameter adjustment of the sequence event construction process. The red dashed line represents the accuracy rate, in which the Alpha parameter values are different aiming at the formula 10. From its trend, when the Alpha value is adjusted from 0.6 to 0.65, and the Beta value is adjusted from 0.35 to 0.4, the quality of analysing the topic evolution is more high and stable, and the accuracy rate is close to 0.70. In general, the parameter adjustment can make the quality of analysing the topic evolution more stable to the maximum for Web news topics, which accords with the experimental effect expected, and can determine the optimal range of the parameter.

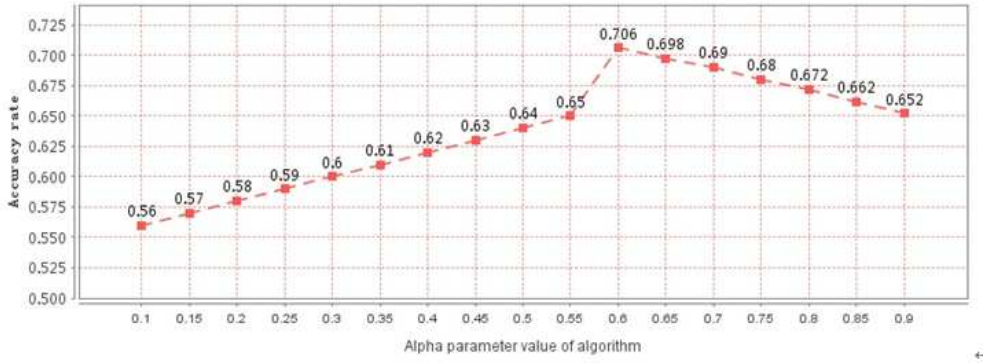


Figure 6: The accuracy rate of analysing the topic evolution with the variation trend for the parameter adjustment

5.6 The accuracy rate analysis with the number of the instances concerned and the similar threshold variation trend

As shown in the figure 7, the accuracy rate represents the quality of analysing the topic evolution by adjusting the similarity threshold of the X axis and the number of the instances of the Y axis. If the threshold is defined, then the quality of analysing the topic evolution is able to increase in a stable trend, because the number of the instances followed is less, the relationship among big data is simpler in the analytical process of the time series and core event similarity.

If the number of the instances followed is gradually increasing, then the relationship among big data adds also the semantic feature for analysing the process of the topic evolution, so its accuracy rate can increase. If the number of the instances is defined, firstly, then the quality of analysing the topic evolution can increase, secondly, then it will decrease with the increasing threshold, because the threshold is less, the inaccurate or approximate accurate analysis of the topic evolution mays be completed. If the threshold can increase to a stable range, then the approximate accurate result of analysing the topic evolution can be excavated. If the threshold can increase to a value, then the accurate result of analysing the topic evolution cannot be excavated. This experiment expresses when the number of the instances followed is one hundred and eighty, and the similarity threshold is zero point seven, the quality of analysing the topic evolution can get the highest about 0.76.

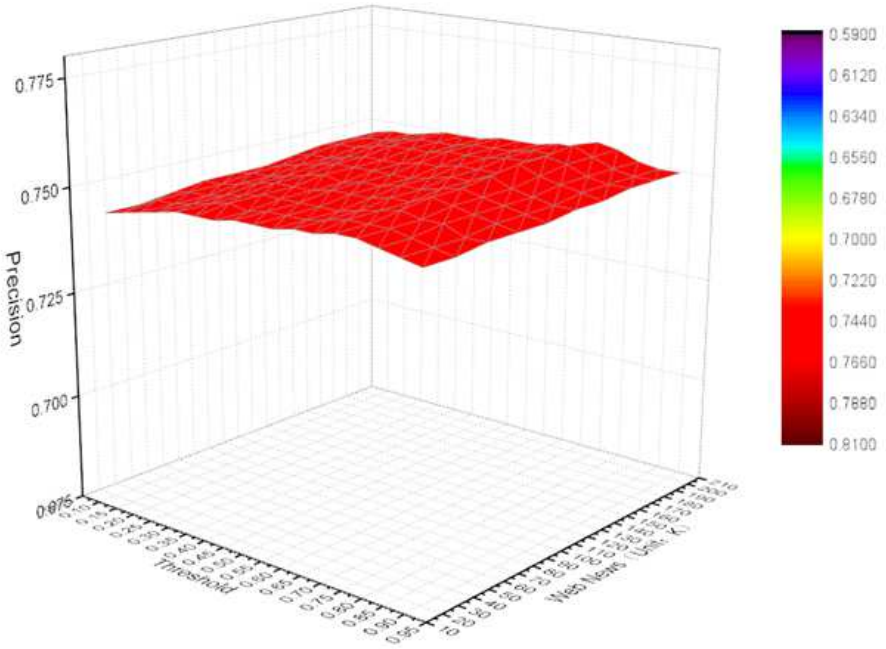


Figure 7: The accuracy rate analysis with the number of the instances concerned and the similar threshold variation trend

5.7 The time consuming analysis for solving main research problem under different methods

As shown in the figure 8, in view of German A320 airliner crash social event, the X axis represents massive Web news time released through the authoritative Web news network platform, the red solid line represents the time consuming for analysing Web hierarchical topic evolution under the method based on the content and description, the blue solid line represents the time consuming for analysing Web hierarchical topic evolution under the method based on the behaviour tracking. According to the change trend of two solid lines, the number of Web news instances has increased sharply in several time intervals with the progress of the event development, therefore, the time consuming has also increased sharply, in other time intervals, the time consuming has relatively stable trend. While the time consuming is lower for the blue solid line, because the analytical process of Web hierarchical topic detection and evolution uses the behaviour tracking method based on the semantic five tuple description and the utility evaluation. This experiment expresses that the non-deterministic problem can be solved efficiently using the

method, which is proposed by this paper.

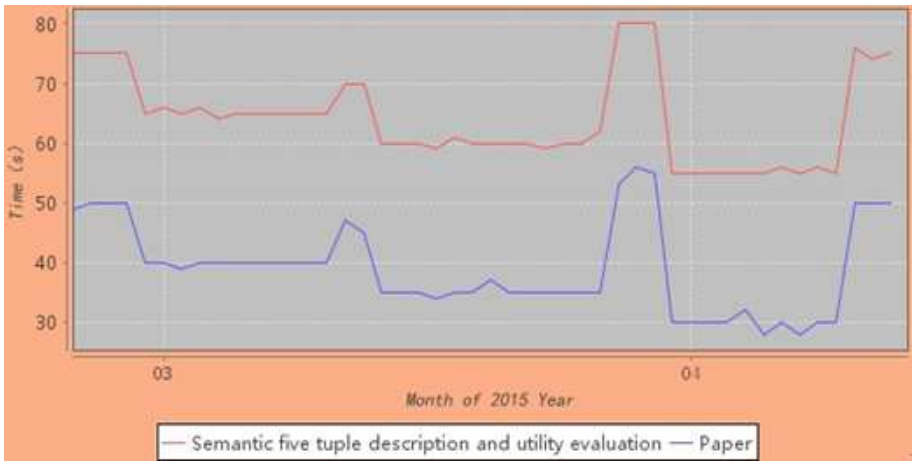


Figure 8: The time consuming analysis for solving main research problem under different methods

5.8 The qualitative analysis of Web hierarchical topic detection and evolution result under different data sets

As shown in the figure 9, the qualitative analysis of Web hierarchical topic detection and evolution result uses also other four standard data sets in addition to the data set of German A320 airliner crash social event, which include Shanghai Bund trample, Taiwan revival airliner falling river, Nepal 8.1 earthquake and Orient Star cruise overturn social event. According to the change trend of five columns, there is little difference in the qualitative analysis of Web hierarchical topic detection and evolution result at the start, development and end stage of five social events. This experiment shows that the analytical process of Web hierarchical topic detection and evolution is stable under different social events. Moreover, the qualitative analysis of Web hierarchical topic detection and evolution result has only little effect in different stages of the social events with the increase of the number of Web news instances.

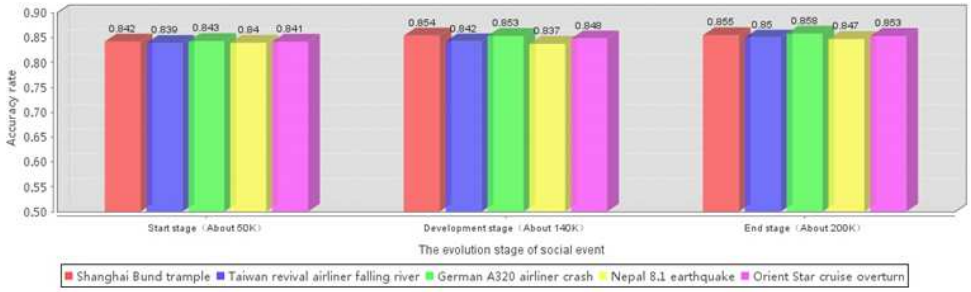


Figure 9: The qualitative analysis of Web hierarchical topic detection and evolution result under different data sets

6 Conclusion

This paper completes the research on an idea of analysing the process for Web hierarchical topic detection and evolution in view of the behaviour tracking technology taking the network

big data of Web news as the processing object, this result of designing and implement is more valuable for the scholars in the research field. In the research process, this paper proposes the analytical algorithm of Web hierarchical topic detection and evolution, in which this paper has also proposed the methods of analysing the usage mode, the topic series construction, the event series construction and the evolution analysis for the events, so as to solve the problem existing in current research status. The results of experimental analysis show that the idea is feasible, verifiable and superior, which plays a major role in reconfiguring Web hierarchical topic corpus, improves understanding efficiency of the network big data, enhances the website availability, constructs and improves the website service function, improves the efficiency of the business operational and website clicking rate, provides an intelligent big data warehouse for the network information evolution application.

Funding

This paper is supported by National Natural Science Foundation of China under Grant Nos.71572015, Support Project of High-Level Teachers in Beijing Municipal Universities in the Period of 13th Five-Year Plan under Grant Nos.CIT&TCD201704072, Premium Funding Project for Academic Human Resources Development in Beijing Union University under Grant Nos.BPHR2018AS01.

Bibliography

[1] Ahila, S.S.; Shunmuganathan, K.L. (2016). Role of Agent Technology in Web Usage Mining: Homomorphic Encryption Based Recommendation for E–commerce Applications, *Wireless Personal Communications*, 87(2), 499-512, 2016.

[2] Alam, M.H.; Ryu, W.J.; Lee, S. (2017). Hashtag-Based Topic Evolution in Social Media, *World Wide Web-Internet and Web Information Systems*, 20(6), 1527-1549, 2017.

[3] Aujla, G.S.; Kumar, N.; Zomaya, A.Y. (2018). Optimal Decision Making for Big Data Processing at Edge-Cloud Environment: An SDN Perspective, *IEEE Transactions on Industrial Informatics*, 14(2), 778–782, 2018.

[4] Chen, B.T.; Tsutsui, S.; Ding, Y.; Ma, F.C. (2017). Understanding the Topic Evolution in a Scientific Domain: an Exploratory Study for the Field of Information Retrieval, *Journal of Informetrics*, 11(4), 1175-1189, 2017.

[5] Chen, M.; Yang, X.P. (2016). Research on Model of Network Information Extraction Based on Improved Topic-Focused Web Crawler Key Technology, *Tehnicki vjesnik/Technical Gazette*, 23(4), 49–54, 2016.

[6] Chen, M.; Yang, X.P.; Sun, M.; Zhao, Y. (2014). Research on Model of Network Information Currency Evaluation Based on Web Semantic Extraction Method, *International Journal of Future Generation Communication and Networking*, 7(2), 103-116, 2014.

[7] Chen, Y.; Zhang, H.; Liu, R.; Ye, Z.W.; Lin, J.Y. (2019). Experimental Explorations on Short Text Topic Mining Between LDA and NMF Based Schemes, *Knowledge-Based Systems*, 163, 1–3, 2019.

[8] Dai, Y.; Wu, W.; Zhou, H.B.; Zhang, J.; Ma, F.Y. (2018). Numerical Simulation and Optimization of Oil Jet Lubrication for Rotorcraft Meshing Gears, *International Journal of Simulation Modelling*, 17(2), 318-326, 2018.

[9] Dai, Y.; Zhu, X.; Zhou, H.; Mao, Z.; Wu, W. (2018). Trajectory Tracking Control for Seafloor Tracked Vehicle by Adaptive Neural-Fuzzy Inference System Algorithm, *International Journal of Computers Communications & Control*, 13(4), 465-476, 2018.

[10] Du, J.; Sun, Y.; Ren, H. (2018). The Relationship of Delivery Frequency with the Cost and Resource Operational Efficiency: A Case Study of Jingdong Logistics, *Mathematics and Computer Science*, 3(6), 129-140, 2018.

[11] Fatima, B.; Ramzan, H.; Asghar, S. (2016). Session Identification Techniques Used in Web Usage Mining a Systematic Mapping of Scholarly Literature, *Online Information Review*, 40(7), 1033-1053, 2016.

[12] Gaul, W.G.; Vincent, D. (2017). Evaluation of the Evolution of Relationships between Topics over Time, *Advances in Data Analysis and Classification*, 11(1), 159-178, 2017.

[13] Jimenez-Marquez, J.L.; Gonzalez-Carrasco, I.; Lopez-Cuadrado, J.L.; Ruiz-Mezcua, B. (2019). Towards a Big Data Framework for Analysing Social Media Content, *International Journal of Information Management*, 44, 1–3, 2019.

[14] Kaseb, M.R.; Khafagy, M.H.; Ali, I.A.; Saad, E.M. (2019). An Improved Technique for Increasing Availability in Big Data Replication, *Future Generation Computer Systems-The International Journal of Escience*, 91, 493–497, 2019.

[15] Kausel, E.E. (2018). Big Data at Work: The Data Science Revolution and Organizational Psychology, *Personnel Psychology*, 71(1), 135-136, 2018.

[16] Kho, N.D. (2018). The State of Big Data, *Econtent*, 41(1), 11-12, 2018.

[17] Liu, J.; Fang, C.; Ansari, N. (2016). Request Dependency Graph: a Model for Web Usage Mining in Large-Scale Web of Things, *IEEE Internet of Things Journal*, 3(4), 598-608, 2016.

[18] Makkie, M.; Huang, H.; Zhao, Y.; Vasilakos, A.V.; Liu, T.M. (2019). Fast and Scalable Distributed Deep Convolutional Autoencoder for fMRI Big Data Analytics, *Neurocomputing*, 325, 20–22, 2019.

[19] Osman, A.M.S. (2019). A Novel Big Data Analytics Framework for Smart Cities, *Future Generation Computer Systems-The International Journal of Escience*, 91, 620–623, 2019.

[20] O'Halloran, K.L.; Tan, S.; Duc-Son, P. (2018). A Digital Mixed Methods Research Design: Integrating Multimodal Analysis with Data Mining and Information Visualization for Big Data Analytics, *Journal of Mixed Methods Research*, 12(1), 11-15, 2018.

[21] Pandian, P.S.; Srinivasan, S. (2016). A Unified Model for Preprocessing and Clustering Technique for Web Usage Mining, *Journal of Multiple-Valued Logic and Soft Computing*, 26(3), 205-220, 2016.

[22] Sagi, T.; Gal, A. (2018). Non-Binary Evaluation Measures for Big Data Integration, *VLDB Journal*, 27(1), 105-110, 2018.

[23] Tran, Q.T.; Nguyen, S.D.; Seo, T.I. (2019). Algorithm for Estimating Online Bearing Fault Upon the Ability to Extract Meaningful Information From Big Data of Intelligent Structures, *IEEE Transactions on Industrial Electronics*, 66(5), 3804–3806, 2019.

[24] Uma, R.; Muneeswaran, K. (2017). OMIR: Ontology-Based Multimedia Information Retrieval System for Web Usage Mining, *Cybernetics and Systems*, 48(4), 393-414, 2017.

[25] Wu, P.J.; Lin, K.C. (2018); Unstructured Big Data Analytics for Retrieving E-Commerce Logistics Knowledge, *Telematics and Informatics*, 35(1), 237-241, 2018.

[26] Yao, L.; Ge, Z.Q. (2019). Scalable Semisupervised GMM for Big Data Quality Prediction in Multimode Processes, *IEEE Transactions on Industrial Electronics*, 66(5), 3681–3684, 2019.

[27] Zhang, D. (2017). High-Speed Train Control System Big Data Analysis Based on Fuzzy RDF Model and Uncertain Reasoning, *International Journal of Computers Communications & Control*, 12(4), 11-15, 2017.

[28] Zhang, D.; Sui, J.; Gong, Y. (2017). Large Scale Software Test Data Generation Based on Collective Constraint and Weighted Combination Method, *Tehnicki Vjesnik*, 24(4), 1041-1050, 2017.

[29] Zhang, D.; Jin, D.; Gong, Y. (2015). Research of Alarm Correlations Based on Static Defect Detection, *Tehnicki vjesnik*, 22(2), 311-318, 2015.

[30] Zhou, H.K.; Yu, H.M.; Hu, R. (2017). Topic Discovery and Evolution in Scientific Literature Based on Content and Citations, *Frontiers of Information Technology & Electronic Engineering*, 18(10), 1511-1524, 2017.

[31] Zhou, H.K.; Yu, H.M.; Hu, R. (2017). Topic Evolution Based on the Probabilistic Topic Model: a Review, *Frontiers of Computer Science*, 11(5), 786-802, 2017.

INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL
ISSN 1841-9836, e-ISSN 1841-9844, 14(3), 329-343, June 2019.

Combination of Evidential Sensor Reports with Distance Function and Belief Entropy in Fault Diagnosis

Y. Dong, J. Zhang, Z. Li, Y. Hu, Y. Deng

Yukun Dong
Institute of Fundamental and Frontier Science
University of Electronic Science and Technology of China, Chengdu, 610054, China

Jiantao Zhang
College of Information Science and Technology
Jinan University, Tianhe, Guangzhou, 510632, China

Zhen Li
College of Information Science and Technology
Jinan University, Tianhe, Guangzhou, 510632, China

Yong Hu
Big Data Decision Institute
Jinan University, Tianhe, Guangzhou, 510632, China

Yong Deng*
1. Institute of Fundamental and Frontier Science
University of Electronic Science and Technology of China, Chengdu, 610054, China
2. Big Data Decision Institute
Jinan University, Tianhe, Guangzhou, 510632, China
*Corresponding author: dengentropy@uestc.edu.cn; prof.deng@hotmail.com

Abstract: Although evidence theory has been applied in sensor data fusion, it will have unreasonable results when handling highly conflicting sensor reports. To address the issue, an improved fusing method with evidence distance and belief entropy is proposed. Generally, the goal is to obtain the appropriate weights assigning to different reports. Specifically, the distribution difference between two sensor reports is measured by belief entropy. The diversity degree is presented by the combination of evidence distance and the distribution difference. Then, the weight of each sensor report is determined based on the proposed diversity degree. Finally, we can use Dempster combination rule to make the decision. A real application in fault diagnosis and an example show the efficiency of the proposed method. Compared with the existing methods, the method not only has a better performance of convergence, but also less uncertainty.
Keywords: Dempster-Shafer evidence theory, sensor data fusion, fault diagnosis, evidence distance, belief entropy, information volume.

1 Introduction

In mechanical engineering, some systems are very complex, which might have many components, reflecting with each other [11,38,44,71]. It is likely that something happens unexpectedly in the systems and causes serious problems due to a variety of reasons, such as unfavorable weather, bad environment or a long time of working. As a result, making full use of sensor reports information is extremely significant to make a reasonable decision in fault diagnosis [58,81].

In order to make a rational decision when using sensor data fusion technology, some works have been proposed to handle uncertainty [39, 45, 59, 77], such as fuzzy set theory [15, 63, 66, 74, 76, 83], Z numbers [30, 31], D numbers [8, 9, 41, 64, 65], R numbers [47, 48] and so on. One

of the most used math tools in sensor data fusion is evidence theory [4, 49]. This theory can efficiently model uncertain information with basic probability assignment (BPA), or called as belief function [21]. In addition, the Dempster rule can efficiently combine the sensor reports from different sources [79]. Due to the desirable properties, evidence theory has been accepted as de facto standard in decision making [10, 28, 78], risk and reliability analysis [13, 27, 40], system optimization [67] and pattern recognition [23, 24, 34, 46, 57].

However, an open issue of evidential sensor data fusion is that the illogical results will be obtained when sensor reports conflict with each other in a high degree [20, 56, 75]. Many methods were presented to address this issue [60, 80]. For example, Yager [70] removed the process of normalizing in D-S combination rule, Smets [50, 51] proposed the conjunctive and disjunctive rules, Murphy [42] combined the conflicting evidence with average operation, Fan and Zuo [16] presented a combination method to fuse conflicting evidence in fault diagnosis, Dubois and Prades method [12], Lefevre *et al.* [33] and so on. Recently, Jiang *et al.* [25] applied belief entropy into sensor data fusion and received the best performance.

Although these methods have some advantages, one of the disadvantages is that some information is not fully used. For example, the distance information and the difference of information volume are not considered in [25]. In this paper, we proposed an improved evidential method, which is conceptually simple, and yet is able to provide much better accuracy and less uncertainty. The basic idea is to obtain the appropriate weights for different reports. The distribution difference between two BPAs is first measured by belief entropy [6]. Then the diversity degree among BPAs can be obtained by combining distribution difference and evidence distance [29]. According to it, the weight of each BPA can be determined. Finally, we can make a decision for fault diagnosis by using Dempster combination rule. An application in fault diagnosis and an example show that our proposed approach can not only increase the accuracy of fault diagnosis but also decrease the uncertain information volume, which is more reasonable.

The remainder of this paper is organized as follows. Section 2 introduces some backgrounds, including Dempster-Shafer Evidence Theory, Evidence Distance and Belief entropy. Section 3 formulates the proposed method for evidential sensor data fusion. A real application in fault diagnosis and an example are given in Section 4 to show the efficiency of the method. The conclusions are in Section 5.

2 Preliminaries

2.1 Dempster-Shafer Evidence Theory

The application in data fusion needs efficient math tools [38]. Dempster-Shafer Evidence Theory, proposed by Dempster [4] and Shafer [49], is effective to handle uncertain information.

Definition 1. Let X be a set of mutually exclusive and collectively exhaustive events, shown as follow [4, 49]:

$$X = \{\theta_1, \theta_2, \cdots, \theta_i, \cdots, \theta_{|X|}\} \tag{1}$$

where set X is called a frame of discernment, whose power set is:

$$2^X = \{\emptyset, \{\theta_1\}, \cdots, \{\theta_{|X|}\}, \{\theta_1, \theta_2\}, \cdots, \{\theta_1, \theta_2, \cdots, \theta_i\}, \cdots, X\} \tag{2}$$

Definition 2. For a frame of discernment $X = \{\theta_1, \theta_2, \cdots, \theta_{|X|}\}$, a mass function is a mapping m from 2^θ to [0,1].

$$m : 2^\theta \to [0, 1] \tag{3}$$

which satisfies the following condition:

$$m(\emptyset) = 0, \sum_{A \in 2^X} m(A) = 1 \tag{4}$$

A, a member of the power set, is called a focal element of the mass function, or named as basic probability assignment (BPA).

BPA is the key issue in evidence theory and many relative processing are presented such as negation [19, 73], correlation [26] and divergence measure [17, 52].

Definition 3. Given two BPAs, m_1 and m_2, they can be combined by,

$$m(A) = \begin{cases} 0, & A = \emptyset \\ \frac{1}{1-K} \sum\limits_{B \cap C = A} m_1(B) m_2(C), & A \neq \emptyset \end{cases} \tag{5}$$

with

$$K = \sum_{B \cap C = \emptyset} m_1(B) m_2(C) \tag{6}$$

where K is a parameter that reflects the conflict between m_1 and m_2. If $K = 0$, m_1 and m_2 have no contradiction.

If $K = 1$, they are totally conflict. Many open issues about conflict management are still not well addressed [2]. Some alternatives are proposed to modify the combination rule [53,54,61,68], others are presented to modify the data models [80] or handle this problem under open world assumption [55, 56].

2.2 Evidence distance

Definition 4. Let m_1 and m_2 be two BPAs on the same frame of discernment of X, which contains N mutually exclusive and exhaustive hypotheses. The distance between m_1 and m_2 is [29]:

$$d_{BPA}(m_1, m_2) = \sqrt{\frac{1}{2}(\overrightarrow{m_1} - \overrightarrow{m_2})^T D(\overrightarrow{m_1} - \overrightarrow{m_2})} \tag{7}$$

where D is a $2^N \times 2^N$ matrix whose elements are

$$D(A, B) = \frac{|A \cap B|}{|A \cup B|}, \quad A, B \in P(X).$$

2.3 Belief entropy

It should be pointed out that uncertainty measurement, decision making and optimization under uncertainty is still an open issue [14,22,37]. Entropy is an efficient tool to model uncertainty [7, 32, 72]. Recently, a new belief entropy, named as Deng entropy was proposed [6]. It has a good performance in measuring uncertainty. Also, it has a backward compatibility, which means when the uncertain information is represented by probability distribution, belief entropy will degenerate to Shannon entropy [1, 3, 35, 43].

Definition 5. Let A be a proposition of BPA, $|A|$ is the cardinality of A. Then, belief entropy is defined as [6]:

$$E_d = -\sum_{A \subseteq X} m(A) log \frac{m(A)}{2^{|A|} - 1} \tag{8}$$

When the BPA has only one element, which means $|A| = 1$, then it can be written as Shannon entropy [6].

$$E_d = -\sum_{A \subseteq X} m(A) log \frac{m(A)}{2^{|A|} - 1} = -\sum_{A \subseteq X} m(A) log\ m(A) \tag{9}$$

3 Evidential sensor data fusion

This section formulates a new weighted average approach for evidential sensor data fusion. The proposed method considered both evidence distance and belief entropy to obtain the appropriate weights assigning to different data reports. Using the weight to pre-treat the multi-source reports, and making the final decision by Dempster combination rule. The flow chart is shown in Figure 1.

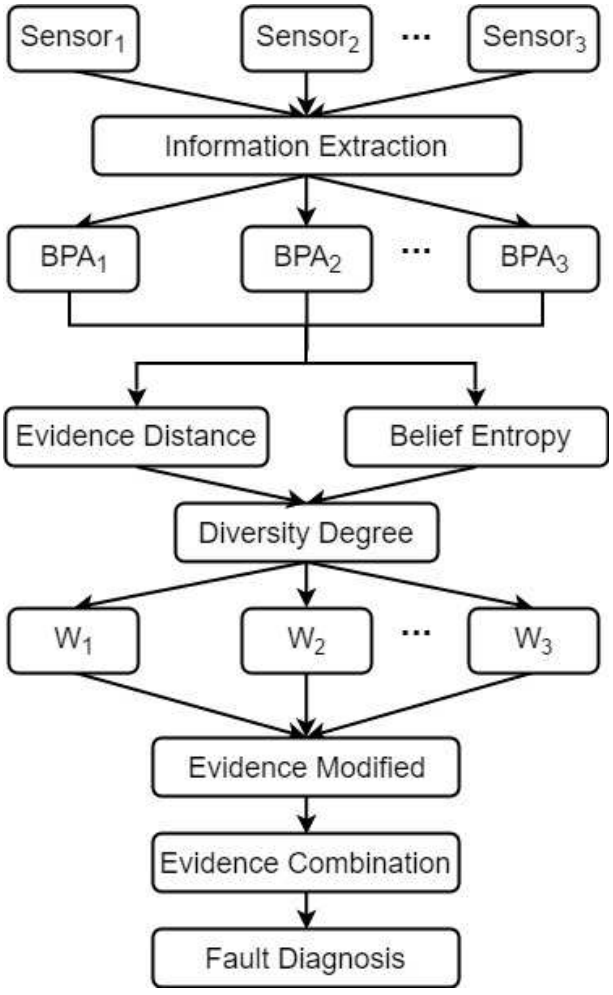


Figure 1: Structure of the proposed sensor data fusing method

Definition 6. Let E_{di}, E_{dj} be the belief entropy of m_1, m_2, then the distribution difference is defined:

$$\alpha_{ij} = e^{|E_{di} - E_{dj}|} \tag{10}$$

Definition 7. Let $d_{BPA}(m_i, m_j)$ be the evidence distance of m_1, m_2, the diversity degree between

two BPAs is defined:

$$D_{ij} = \frac{1}{\alpha_{ij} \times d_{BPA}(m_i, m_j)} \tag{11}$$

Definition 8. A diversity of distribution and distance matrix (DDD) is also defined:

$$DDD = \begin{pmatrix} 1 & D_{12} & \cdots & D_{1i} & \cdots & D_{1n} \\ D_{21} & 1 & \cdots & D_{2i} & \cdots & D_{2n} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ D_{i1} & D_{i2} & \cdots & 1 & \cdots & D_{in} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ D_{n1} & D_{n2} & \cdots & D_{ni} & \cdots & 1 \end{pmatrix} \tag{12}$$

The computation step is shown as following.

- Step 1, collect sensor information and transform into BPAs.
- Step 2, use Equation (7) to calculate the evidence distance between two BPAs, which shows the difference.
- Step 3, use Equation (8) to calculate the belief entropy of BPA, which shows the distributive characteristic.
- Step 4, calculate the diversity degree between two BPAs.
- Step 5, the support of the BPAs is given as:

$$Sup(m_i) = \sum_{j=1,\ j \neq i}^{n} D_{ij} \tag{13}$$

- Step 6, the credibility degree of the BPAs is obtained.

$$Crd_i = \frac{Sup(m_i)}{\sum_{i=1}^{n} Sup(m_i)} \tag{14}$$

- Step 7, it is easy to see that $\sum_{i=1}^{n} Crd_i = 1$. As a result, Crd_i can be the weight of each BPA. $(\omega_1, \omega_2, \cdots, \omega_n) = (Crd_1, Crd_2, \cdots, Crd_n)$. Thus, a new weighted evidence can be obtained, which is:

$$m(a) = \omega_1 \times m_1(a) + \omega_2 \times m_2(a) + \cdots + \omega_n \times m_n(a) \tag{15}$$

- Step 8, use Dempster combination rule to combine the new weighted evidence to get the result. Furthermore, if the number of original evidence is n, then the new evidence should be combined for $(n-1)$ times [25].

It should be noticed that if $d_{BPA}(m_i, m_j) = 0$, which means that there is no conflict between m_i and m_j. In this situation, the value of diversity degree D_{ij} cannot be determined. So we proposed our own solution, which can be discussed further.

First, if there are only three BPAs m_1, m_2, m_3 and $m_1 = m_2$, the two BPAs can be regarded as the same one, then use Murphy's method, which is assigning the weight equally to each BPA.

Second, if there are more than three evidences. Supposing there are n BPAs, which can be divided into m groups according to their d_{BPA}. If there are only two groups, it will just be the same as the first situation. If there are more than two groups, and each group has k BPAs, noted as $k_1, k_2, \cdots, k_m$, and $\sum_{i=1}^{m} k_i = n$.

$$\begin{array}{cccc} \{m_{11} & m_{12} & \cdots & m_{1k_1}\} \\ \{m_{21} & m_{22} & \cdots & m_{2k_2}\} \\ \vdots & \vdots & \vdots & \vdots \\ \{m_{m1} & m_{12} & \cdots & m_{mk_m}\} \end{array} \tag{16}$$

Then, select one BPA from each group. Suppose they are $m_{11}, m_{21}, \cdots, m_{m1}$, and use the proposed method to obtain the weight, which is

$$\{w(m_{11}), w(m_{21}), \cdots, w(m_{m1})\} = \{\omega_1, \omega_2, \cdots, \omega_m\} \tag{17}$$

Since $w = \sum_{i=1}^{m} k_i \omega_i$, and the finally weight can be gained.

$$\begin{aligned} &\{\omega_{11}, \omega_{12}, \cdots, \omega_{1k_1}, \omega_{21}, \omega_{22}, \cdots, \omega_{2k_2}, \cdots, \omega_{m1}, \omega_{m2}, \cdots, \omega_{mk_m}\} \\ &= \{\underbrace{\frac{\omega_1}{w}, \frac{\omega_1}{w}, \cdots, \frac{\omega_1}{w}}_{k_1}, \underbrace{\frac{\omega_2}{w}, \frac{\omega_2}{w}, \cdots, \frac{\omega_2}{w}}_{k_2}, \cdots, \underbrace{\frac{\omega_m}{w}, \frac{\omega_m}{w}, \cdots, \frac{\omega_m}{w}}_{k_m}\} \end{aligned} \tag{18}$$

Two examples are given to illustrate how it works.

Example 1. Suppose the frame of discernment is $X = \{a_1,\ a_2,\ a_3\}$, and there are three BPAs.

$$m_1(a_1) = 0.3,\ m_1(a_2, a_3) = 0.7$$

$$m_2(a_1) = 0.3,\ m_2(a_2, a_3) = 0.7$$

$$m_3(a_1) = 0.8,\ m_3(a_2, a_3) = 0.2$$

Then the weight should be obtained as $W(1/4, 1/4, 1/2)$.

Example 2. Suppose the frame of discernment is also $X = \{a_1,\ a_2,\ a_3\}$, and there are four BPAs.

$$m_1(a_1) = 0.3,\ m_1(a_2, a_3) = 0.7,$$

$$m_2(a_1) = 0.3,\ m_2(a_2, a_3) = 0.7,$$

$$m_3(a_1) = 0.8,\ m_3(a_2, a_3) = 0.2.$$

$$m_4(a_1) = 0.6,\ m_4(a_2, a_3) = 0.4.$$

In this example, we can just consider m_2, m_3 and m_4, then use the proposed method to gain the weights, which are

$$\omega_2 = 0.2585, \omega_3 = 0.3072, \omega_4 = 0.4343$$

Finally, re-assign the weights to get the result: $W = (0.2054, 0.2054, 0.2441, 0.3451)$.

4 Experiments

4.1 Application in fault diagnosis

The complex systems is very complicated since each factor in the system interacting with each other in a very complicated way. To address this issue, network analysis [18,36,62,69,82] and data fusion based technology are presented to deal with complexity and guarantee the reliability of the complex system. [25] gave a case of motor rotor fault diagnosis, where the vibration signal is collected by acceleration sensor (m_1), velocity sensor (m_2), and displacement sensor (m_3). The possible faults including normal operation (F_1), unbalance (F_2), misalignment (F_3), pedestal looseness (F_4). The collected data report is shown in Table 1, where X is a frame of discernment, and $X = \{F_1, F_2, F_3, F_4\}$.

Table 1: Output of the multi-senors [25]

m_i	F_1	F_2	F_3	F_4	X
m_1	0.06	0.68	0.02	0.04	0.20
m_2	0.02	0	0.79	0.05	0.14
m_3	0.02	0.58	0.16	0.04	0.20

The computation steps are as following.
By Equation (7), the evidence distance between two BPAs,

$$d_{BPA}(m_1, m_2) = 0.7276, \quad d_{BPA}(m_1, m_3) = 0.1249, \quad d_{BPA}(m_2, m_3) = 0.6063$$

Using Equation (8) to get the belief entropy of each BPA.

$$E_{d1} = 2.1663, \quad E_{d2} = 1.5417, \quad E_{d3} = 2.4232$$

The distribution difference is obtained by Equation (10).

$$\alpha_{12} = 1.8674, \quad \alpha_{13} = 1.2930, \quad \alpha_{23} = 2.4145$$

And the diversity degree calculated by Equation (11) are:

$$D(m_1, m_2) = 0.7360, \quad D(m_1, m_3) = 6.1923, \quad D(m_2, m_3) = 0.6831$$

Finally, the weights by Equation (14):

$$\omega_1 = 0.4551, \quad \omega_2 = 0.0932, \quad \omega_3 = 0.4517$$

According to the weights, a new set of data can be got by calculating:

$$m(F_1) = 0.06 \times 0.4551 + 0.02 \times 0.0932 + 0.02 \times 0.4517 = 0.0382$$

In the same way, we can calculate $m(F_2) = 0.5714, m(F_3) = 0.1550, m(F_4) = 0.0409, m(X) = 0.1944$.

Therefore, the decision can be made by fusing $\{0.0382, 0.5714, 0.1550, 0.0409, 0.1944\}$ with Dempster combination rule for 2 times, which are shown in Table 2. The results of other methods are also listed as a comparison. Following [25], we will use $\Delta = 0.7$ as the threshold.

Since the fault F_2 has a belief degree of 89.18%, it can be told that unbalance is the fault of the equipment. And compared with other methods, the new method performs much better. The belief entropy in Figure 2 is the smallest, which means that the proposed method has the smallest uncertain information volume.

Table 2: Comparison of the result of several existing methods

Combination rule	F_1	F_2	F_3	F_4	X	Diagnosis result
Dempster [4, 49]	0.0205	0.5229	0.3933	0.0309	0.0324	Uncertainty
Murphy [42]	0.0112	0.6059	0.3508	0.0153	0.0168	Uncertainty
Jiang [25]	0.0111	0.7265	0.2313	0.0144	0.0168	Unbalance
Deng [5]	0.0111	0.7728	0.1851	0.0139	0.0165	Unbalance
Proposed method	0.0106	0.8918	0.0713	0.0115	0.0148	Unbalance

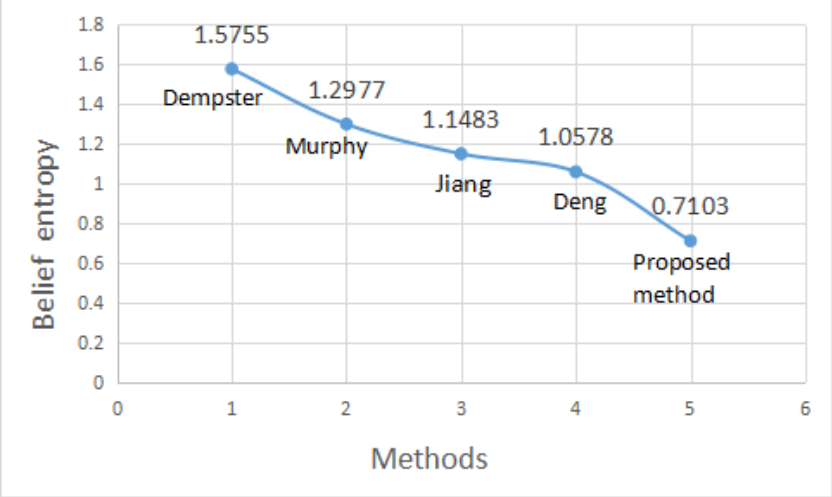


Figure 2: Belief entropy comparison.

4.2 Example

The proposed method is not only efficient in dealing with uncertainty, but also be false-evidence resilient. Suppose the system has collected five evidences from five different sensors [5], which are shown as follows:

$$
\begin{aligned}
m_1 &: m_1(A) = 0.5,\ m_1(B) = 0.2,\ m_1(C) = 0.3;\\
m_2 &: m_2(A) = 0,\ m_2(B) = 0.9,\ m_2(C) = 0.1;\\
m_3 &: m_3(A) = 0.55,\ m_3(B) = 0.1,\ m_3(C) = 0.35;\\
m_4 &: m_4(A) = 0.55,\ m_4(B) = 0.1,\ m_4(C) = 0.35;\\
m_5 &: m_5(A) = 0.55,\ m_5(B) = 0.1,\ m_5(C) = 0.35;
\end{aligned}
$$

The results are shown in Table 3, Obviously, the second evidence m_2 is conflicting with the others. Although there are more and more evidences that support A, Dempster's method cannot reach a reasonable result. $m(A)$ will always be zero as long as one evidence does not support it. While other methods revise the disadvantage. Apparently, the proposed method can have a probability 0.7663 to support A when the third evidence is fused, and it always has the highest support.

In previous work, some information is not fully used. In our proposed method, not only the distance information between BPA, but also the difference of information volume of each BPA is considered. As a result, our method can efficiently measure the support degree of each sensor report. That is, if one report is more reliable, more weights will be assigned to this report. Therefore, the result of our improved method is more reasonable and more desirable.

Despite the advantages, the proposed method might not be very suitable in some situations.

For example, the calculation process will be extremely complex if the data of evidence is huge. [25] considered belief entropy, and considered evidence distance, while our method considered both of them. Therefore, the calculation complexity might be twice as those methods. In addition, the proposed method might reach an unreasonable result if a bad evidence repeats many times. This is because the bad evidence cannot be identified, and they could have a mutual confirmation, leading to an illogical fusing result.

Table 3: Results of different method

Combination rule	$m_1,\ m_2$	$m_1,\ m_2,\ m_3$	$m_1,\ m_2,\ m_3,\ m_4$	$m_1,\ m_2,\ m_3,\ m_4,\ m_5$
Dempster [4, 49]	m(A)=0	m(A)=0	m(A)=0	m(A)=0
	m(B)=0.8571	m(B)=0.6316	m(B)=0.3288	m(B)=0.1228
	m(C)=0.1429	m(C)=0.3684	m(C)=0.6712	m(C)=0.8772
Murphy [42]	m(A)=0.1543	m(A)=0.3500	m(A)=0.6027	m(A)=0.7958
	m(B)=0.7469	m(B)=0.5524	m(B)=0.2627	m(B)=0.0932
	m(C)=0.0988	m(C)=0.1276	m(C)=0.1346	m(C)=0.1110
Jiang [25]	m(A)=0.4206	m(A)=0.6819	m(A)=0.8244	m(A)=0.8945
	m(B)=0.3944	m(B)=0.1310	m(B)=0.0326	m(B)=0.0071
	m(C)=0.1850	m(C)=0.1871	m(C)=0.1430	m(C)=0.0984
Deng [5]	m(A)=0.1543	m(A)=0.5816	m(A)=0.8061	m(A)=0.8909
	m(B)=0.7469	m(B)=0.2439	m(B)=0.0481	m(B)=0.0086
	m(C)=0.0988	m(C)=0.1745	m(C)=0.1457	m(C)=0.1005
Proposed method	m(A)=0.1543	m(A)=0.7663	m(A)=0.8554	m(A)=0.9063
	m(B)=0.7469	m(B)=0.0424	m(B)=0.0082	m(B)=0.0015
	m(C)=0.0988	m(C)=0.1913	m(C)=0.1364	m(C)=0.0922

5 Conclusions

In fault diagnosis and other sensor data fusion systems, the reports of different sensors may be influenced by some complex environments, leading them less reliable. Therefore, how to efficiently determine the reliability of each report, or to say, the weight of each report is very important. To address this issue, we propose an improved method based on the belief entropy and the evidence distance. The method considers both the degree of conflict and the difference of information volume among evidences. An application and an example illustrate the efficiency of the method in evidential sensor fusion. It shows that the proposed method is more efficient for highly conflicting evidences with better performance of convergence and less uncertainty, compared with the existing methods.

Some related advantages and disadvantages of different methods are discussed as follows.

- Dempster's method, which can well deal with imprecise and uncertain information, is widely used in fusing information. However, when it comes to highly conflicting evidences, the method will always lead to some illogical results.

- Murphy's method, which uses an average operation to combine the conflicting evidence, is able to deal with highly conflicting evidences to some extent. However, the difference and relationship of evidences is neglected.

- In Jiang's method, belief entropy is used to calculate the weight of each evidence. It considers the difference of evidence, which makes it more reasonable than Murphy's method.

- In Deng's method, which is different from Jiang's method, evidence distance is used to calculate the weight rather than the belief entropy. And the similarity between two evidences is proposed.
- The proposed method, considering both belief entropy and evidence distance, has a better result. Although it might be not very suitable in some situations, it leverages the advantages of Jiang's method and Deng's method.

Funding

The work is partially supported by National Natural Science Foundation of China (Grant Nos. 61573290, 61503237).

Author contributions. Conflict of interest

The authors contributed equally to this work. The authors declare no conflict of interest.

Bibliography

[1] Abellan, J. (2017). Analyzing properties of Deng entropy in the theory of evidence, *Chaos Solitons & Fractals*, 95, 195–199, 2017.

[2] An, J.Y.; Hu, M.; Fu, L.; Zhan, J.W. (2019). A novel fuzzy approach for combining uncertain conflict evidences in the Dempster-Shafer theory, *IEEE Access*, 7, 7481–7501, 2019.

[3] Cui, H.; Liu, Q.; Zhang, J.; Kang, B. (2019). An improved Deng entropy and its application in pattern recognition, *IEEE Access*, 7,18284–18292, 2019.

[4] Dempster, A.P. (1967). Upper and lower probabilities induced by a multivalued mapping, *Annals of Mathematics and Statistics*, 38(2), 325–339, 1967.

[5] Deng, Y.; Shi, W.K.; Zhu, Z.F.; Liu, Q. (2004). Combining belief functions based on distance of evidence, *Decision Support Systems*, 38(3), 489–493, 2004.

[6] Deng, Y. (2106). Deng entropy, *Chaos, Solitons & Fractals*, 91, 549–553, 2016.

[7] Deng, W.; Deng, Y. (2018). Entropic methodology for entanglement measures, *Physica A: Statistical Mechanics and its Applications*, 512, 693–697, 2018.

[8] Deng, X.Y.; Deng, Y. (2019). D-AHP method with different credibility of information, *Soft Computing*, 23(2), 683–691, 2019.

[9] Deng, X.Y.; Jiang, W. (2019). D number theory based game-theoretic framework in adversarial decision making under a fuzzy environment, *International Journal of Approximate Reasoning*, 106, 194–213, 2019.

[10] Deng, X.Y.; Jiang, W.; Wang, Z. (2019). Zero-sum polymatrix games with link uncertainty: A Dempster-Shafer theory solution, *Applied Mathematics and Computation*, 340, 101–112, 2019.

[11] Dong, Y.K.; Wang, J.Y.; Chen, F.H.; Hu, Y.; Deng, Y. (2017). Location of Facility Based on Simulated Annealing and "ZKW" Algorithms, *Mathematical Problems in Engineering*, 2017, 9, 2017.

[12] Dubois, D.; Prade, H. (1988). Representation and combination of uncertainty with belief functions and possibility measures, *Computational Intelligence*, 4(3), 244–264, 1988.

[13] Dutta, P. (2017). Modeling of variability and uncertainty in human health risk assessment, *MethodsX*, 4, 76–85, 2017.

[14] Dutta, P. (2018). An uncertainty measure and fusion rule for conflict evidences of big data via Dempster-Shafer theory, *International Journal of Image and Data Fusion*, 9(2), 152–169, 2018.

[15] Dzitac, I.; Filip, F.G.; Manolescu, M.J. (2017). Fuzzy logic is not fuzzy: World-renowned computer scientist Lotfi A. Zadeh, *International Journal of Computers Communications & Control*, 12(6), 748–789, 2017.

[16] Fan, X.F.; Zuo, M.J. (2006). Fault diagnosis of machines based on D-S evidence theory. Part 1: D-S evidence theory and its improvement, *Pattern Recognition Letters*, 27(5), 366–376, 2006.

[17] Fei, L.G.; Deng, Y. (2019). A new divergence measure for basic probability assignment and its applications in extremely uncertain environments, *International Journal of Intelligent Systems*, 34(4), 584–600, 2019.

[18] Fei, L.; Deng, Y. (2018). Identifying influential nodes in complex networks based on the inverse-square law, *Physica A: Statistical Mechanics and its Applications*, 512, 1044–1059, 2018.

[19] Gao, X.; Deng, Y. (2019). The negation of basic probability assignment, *IEEE Access*, 7, 10.1109/ACCESS.2019.2901932, 2019.

[20] Haenni, R. (2002). Are alternatives to Dempster's rule of combination real alternatives?: Comments on: About the belief function combination and the conflict management problem—-Lefevre et al, *Information Fusion*, 3(3), 237–239, 2002.

[21] Han, Y.; Deng, Y. (2019). A novel matrix game with payoffs of Maxitive Belief Structure, *International Journal of Intelligent Systems*, 34(4), 690–706, 2019.

[22] Han, Y.; Deng, Y. (2018). A hybrid intelligent model for Assessment of critical success factors in high risk emergency system, *Journal of Ambient Intelligence and Humanized Computing*, 9(6), 1933–1953, 2018.

[23] Han, Y.; Deng, Y. (2018). An Evidential Fractal AHP target recognition method, *Defence Science Journal*, 68(4), 367–373, 2018.

[24] Jaunzemis, A.D.; Holzinger, M.J.; Chan, M.W.; Shenoy, P.P. (2019). Evidence gathering for hypothesis resolution using judicial evidential reasoning, *Information Fusion*, 49, 26–45, 2019.

[25] Jiang, W.; Wei, B.Y.; Xie, C.H.; Zhou, D.Y. (2016). An evidential sensor fusion method in fault diagnosis, *Advances in Mechanical Engineering*, 8(3), 1–7, 2016.

[26] Jiang, W. (2018). A correlation coefficient for belief functions, *International Journal of Approximate Reasoning*, 103, 94–106, 2018.

[27] Jiang, W.; Wang, S.Y. (2017). An Uncertainty Measure for Interval-valued Evidences, *International Journal of Computers Communications & Control*, 12(5), 631–644, 2017.

[28] Jin, L.Q.; Fang, X. (2017). Interval Certitude Rule Base Inference Method using the Evidential Reasoning, *International Journal of Computers Communications & Control*, 12(6), 2017.

[29] Jousselme, A.L.; Grenier, D.; Bossé, É, (2001). A new distance between two bodies of evidence, *Information Fusion*, 2(2), 91–101, 2001.

[30] Kang, B.Y.; Zhang, P.D.; Gao, Z.Y.; Chhipi-Shrestha, G.; Hewage, K.; Sadiq, R. (2019). Environmental assessment under uncertainty using Dempster–Shafer theory and Z-numbers, *Journal of Ambient Intelligence and Humanized Computing*, https://doi.org/10.1007/s12652-019-01228-y, 2019.

[31] Kang, B.Y.; Deng, Y.; Hewage, K.; Sadiq, R. (2019). A method of measuring uncertainty for Z-number, *IEEE Transactions on Fuzzy Systems*, 27(4), 731-738, 2019.

[32] Kuzemsky, A.L. (2018). Temporal evolution, directionality of time and irreversibility, *Rivista Del Nuovo Cimento*, 41(10), 513–574, 2018.

[33] Lefevre, E.; Colot, O.; Vannoorenberghe, P. (2002). Belief function combination and conflict management, *Information fusion*, 3(2), 149–162, 2002.

[34] Li, M.Z.; Zhang, Q.; Deng, Y. (2018). Evidential identification of influential nodes in network of networks, *Chaos, Solitons & Fractals*, 117, 283–296, 2018.

[35] Li, Y.X.; Deng, Y. (2018). Generalized Ordered Propositions Fusion Based on Belief Entropy, *International Journal of Computers Communications & Control*, 13(5), 792–807, 2018.

[36] Liu, C.;Li, L.; Wang, Z.; Wang, R.(2019). Pattern transitions in a vegetation system with cross-diffusion, *Applied Mathematics and Computation*, 342, 255–262, 2019.

[37] Liu, H.; Dzitac, I.; Guo, S.(2018). Reduction of Conditional Factors in Causal Analysis, *International Journal of Computers Communications & Control*, 13(3), 383–390, 2018.

[38] Meng, D.; Yang, S.; Zhang, Y.; Zhu, S (2018). Structural reliability analysis and uncertainties-based collaborative design and optimization of turbine blades using surrogate model, *Fatigue & Fracture of Engineering Materials & Structures*, 10.1111/ffe.12906, 2018.

[39] Meng, D.; Liu, M.; Yang, S.; Zhang, H.; Ding, R. (2018). A fluid–structure analysis approach and its application in the uncertainty-based multidisciplinary design and optimization for blades, *Advances in Mechanical Engineering*, 10(6), 1687814018783410, 2018.

[40] Meng, D.; Li, Y.; Zhu, S.; Lv, G.; Correia, J.; de Jesus, A.(2019). An Enhanced Reliability Index Method and Its Application in Reliability-Based Collaborative Design and Optimization, *Mathematical Problems in Engineering*, https://doi.org/10.1155/2019/4536906, 2019.

[41] Mo, H.M.; Deng, Y. (2019). An evaluation for sustainable mobility extended by D numbers, *Technological and Economic Development of Economy*, Accepted, 2019.

[42] Murphy, C.K. (2000); Combining belief functions when evidence conflicts, *Decision Support Systems*, 29(1), 1–9, 2000.

[43] Pan, L.P.; Deng, Y. (2018). A New Belief Entropy to Measure Uncertainty of Basic Probability Assignments Based on Belief Function and Plausibility Function, *Entropy*, 20(11), 842, 2018.

[44] Paté-Cornell, M.E. (1990). Organizational Aspects of Engineering System, *Safety Science*, 250, 1210–16, 1990.

[45] Rong, H.; Ge, M.; Zhang, G.; Zhu, M. (2018). An approach for detecting fault lines in a small current grounding system using fuzzy reasoning spiking neural p systems, *International Journal of Computers Communications & Control*, 13(4), 521–536, 2018.

[46] Sabahi, F. (2016). A Novel Generalized Belief Structure Comprising Unprecisiated Uncertainty Applied to Aphasia Diagnosis, *Journal of Biomedical Informatics*, 62, 66–77, 2016.

[47] Seiti, H.; Hafezalkotob, A. (2019). Developing the R-TOPSIS methodology for risk-based preventive maintenance planning: A case study in rolling mill company, *Computers & Industrial Engineering*, 128, 622–636, 2019.

[48] Seiti, H.; Hafezalkotob, A.; Martinez, L. (2019). R-numbers, a new risk modeling associated with fuzzy numbers and its application to decision making, *Information Sciences*, 483, 206–231, 2019.

[49] Shafer, G. (1967). *A Mathematical Theory of Evidence*, Princeton University Press, 1967.

[50] Smets, P. (1993). Belief functions: the disjunctive rule of combination and the generalized Bayesian theorem, *International Journal of Approximate Reasoning*, 9(1), 1–35, 1993.

[51] Smets, P.; Kennes, R. (1994). The transferable belief model, *Artificial Intelligence*, 66(2), 191–234, 1994.

[52] Song, Y.T.; Deng, Y. (2019). A new method to measure the divergence in evidential sensor data fusion, *International Journal of Distributed Sensor Networks*, 15(4), DOI: 10.1177/1550147719841295, 2019.

[53] Su, X.Y.; Li, L.S.; Shi, F.J.; Qian, H. (2018). Research on the Fusion of Dependent Evidence Based on Mutual Information, *IEEE Access*, 6, 71839–71845, 2018.

[54] Su, X.Y.; Li, L.S.; Qian, H.; Sankaran, M.; Deng, Y. (2019). A new rule to combine dependent bodies of evidence, *Soft Computing*, https://doi.org/10.1007/s00500-019-03804-y, 2019.

[55] Sun, R.L.; Deng, Y. (2019); A new method to identify incomplete frame of discernment in evidence theory, *IEEE Access*, 7(1), 15547–15555, 2019.

[56] Sun, R.L.; Deng, Y. (2019). A new method to determine generalized basic probability assignment in the open world, *IEEE Access*, 7(1), accepted, 2019.

[57] Vandoni, J.; Aldea, E.; Le Hégarat-Mascle, S. (2019). Evidential query-by-committee active learning for pedestrian detection in high-density crowds, *International Journal of Approximate Reasoning*, 104, 166–184, 2019.

[58] Wang, T.; Zhang, G.X.; Rong, H.N.; Pérez-Jiménez, M.J. (2014). Application of fuzzy reasoning spiking neural P systems to fault diagnosis, *International Journal of Computers Communications & Control*, 9(6), 786–799, 2014.

[59] Wang, Y.; Wang, S.; Deng, Y. (2019). A modified efficiency centrality to identify influential nodes in weighted networks, *Pramana*, 68(4), 68, 2019.

[60] Wang, Y.J.; Deng, Y. (2018). Base belief function: an efficient method of conflict management, *Journal of Ambient Intelligence and Humanized Computing*, https://doi.org/10.1007/s12652-018-1099-2, 2018.

[61] Wang, J.; Qiao, K.Y.; Zhang, Z.Y. (2019). An improvement for combination rule in evidence theory, *Future Generation Computer Systems*, 91, 1–9, 2019.

[62] Wei, B.; Deng, Y. (2018). A cluster-growing dimension of complex networks: From the view of node closeness centrality, *Physica A: Statistical Mechanics and its Applications*, 10.1016/j.physa.2019.01.125, 2019.

[63] Xiao, F.Y. (2018). A Hybrid Fuzzy Soft Sets Decision Making Method in Medical Diagnosis, *IEEE Access*, 6, 25300–25312, 2018.

[64] Xiao, F.Y. (2018); A novel multi-criteria decision making method for assessing health-care waste treatment technologies based on D numbers, *Engineering Applications of Artificial Intelligence*, 71(2018), 216–225, 2018.

[65] Xiao, F.Y. (2019). A multiple criteria decision-making method based on D numbers and belief entropy, *International Journal of Fuzzy Systems*, https://doi.org/10.1007/s40815-019-00620-2, 2019.

[66] Xiao, F.Y.; Ding, W.P. (2019). A novel multi-criteria decision making method for assessing health-care waste treatment technologies based on D numbers, *Applied Soft Computing*, 74, DOI: 10.1016/j.asoc.2019.03.043, 2019.

[67] Xu, X.B.; Li, S.B.; Song, X.J.; Wen, C.L.; Xu, D.L. (2016). The optimal design of industrial alarm systems based on evidence theory, *Control Engineering Practice*, 46, 142–156, 2016.

[68] Xu, H.; Deng, Y. (2019). Dependent Evidence Combination Based on DEMATEL Method, *International Journal of Intelligent Systems*, 34, 10.1002/int.22107, 2019.

[69] Xu, P.; Zhang, R.; Deng, Y. (2019). A Novel Visibility Graph Transformation of Time Series into Weighted Networks, *Chaos, Solitons & Fractals*, 2018, 201–208, 2018.

[70] Yager, R.R. (1987). On the Dempster-Shafer framework and new combination rules, *Information Sciences*, 41(2), 93–137, 1987.

[71] Yang, H.C.; Deng, Y.; Jones, J. (2018). Network Division Method Based on Cellular Growth and Physarum-inspired Network Adaptation, *International Journal of Unconventional Computing*, 13(6), 477–491, 2018.

[72] Yin, L.K.; Deng, Y. (2018). Toward uncertainty of weighted networks: An entropy-based model, *Physica A: Statistical Mechanics and its Applications*, 508, 176–186, 2018.

[73] Yin, L.K.; Deng, X.Y.; Deng, Y. (2019). The negation of a basic probability assignment, *IEEE Transactions on Fuzzy Systems*, 27(1), 135–143, 2019.

[74] Zadeh, L.A. (1965). Fuzzy Sets, *Information and Control*, 8(3), 338–353, 1965.

[75] Zadeh, L.A. (1986). A simple view of the Dempster-Shafer theory of evidence and its implication for the rule of combination, *AI magazine*, 7(2), 85, 1986.

[76] Zavadskas, E.K.; Antucheviciene, J.; Hajiagha, S.H.R. (2015). The interval-valued intuitionistic fuzzy MULTIMOORA method for group decision making in engineering, *Mathematical Problems in Engineering*, 2015, 13, 2015.

[77] Zavadskas, E.K.; Antucheviciene, J.; Turskis, Z.; Adeli, H. (2016). Hybrid multiple-criteria decision-making methods: A review of applications in engineering, *Scientia Iranica. Transaction A, Civil Engineering*, 23(1), 1, 2016.

[78] Zhang, L.M.; Wu, X.G.; Qin, Y.W.; Skibniewski, M.J.; Liu, W.L. (2016). Towards a Fuzzy Bayesian Network Based Approach for Safety Risk Analysis of Tunnel-Induced Pipeline Damage, *Risk Analysis*, 36(2), 278–301, 2016.

[79] Zhang, L.M.; Wu, X.G.; Zhu, H.P.; AbouRizk, S.M. (2017). Perceiving safety risk of buildings adjacent to tunneling excavation: An information fusion approach, *Automation in Construction*, 73, 88–101, 2017.

[80] Zhang, W.Q.; Deng, Y. (2018). Combining conflicting evidence using the DEMATEL method, *Soft computing*, https://doi.org/10.1007/s00500-018-3455-8, 2018.

[81] Zhang, H.P.; Deng, Y. (2018). Engine fault diagnosis based on sensor data fusion considering information quality and evidence theory, *Advances in Mechanical Engineering*, 10(11), DOI: 10.1177/1687814018809184, 2018.

[82] Zhao, D.; Wang, L.; Wang, Z.; Xiao, G. (2019). Virus Propagation and Patch Distribution in Multiplex Networks: Modeling, Analysis, and Optimal Allocation, *IEEE Transactions on Information Forensics and Security*, 14(7), 1755–1767, 2019.

[83] Zhu, W.B.; Yang, H.C.; Jin, Y.; Liu, B.Y. (2017). A Method for Recognizing Fatigue Driving Based on Dempster-Shafer Theory and Fuzzy Neural Network, *Mathematical Problems in Engineering*, 10, 2017.

INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL
ISSN 1841-9836, e-ISSN 1841-9844, 14(3), 344-358, June 2019.

Selection of Cluster Heads for Wireless Sensor Network in Ubiquitous Power Internet of Things

W. Hu, W.H. Yao, Y.W. Hu, H.H. Li

Wei Hu, Wenhui Yao*, Yawei Hu, Huanhao Li
School of Economics and Management,
Shanghai University of Electric Power,
Shanghai 200090, China
*Corresponding author: 272828408@qq.com

Abstract: This paper designs a selection algorithm of cluster heads (CHs) in wireless sensor network (WSN) under the ubiquitous power Internet of Things (UPIoT), aiming to solve the network failure caused by premature death of WSN sensors and overcome the imbalance in energy consumption of sensors. The setting of the cluster head node helps to reduce the energy consumption of the nodes in the network, so the choice of cluster head is very important. The author firstly explains the low energy adaptive clustering hierarchy (LEACH) and the distance and energy based advanced LEACH (DEAL) protocol. Compared with the LEACH, the DEAL considers the remaining nodal energy and the sensor-sink distance. On this basis, the selectivity function-based CH selection (SF-CHs) algorithm was put forward to select CHs and optimize the clustering. Specifically, the choice of CHs was optimized by a selectivity function, which was established based on the remaining energy, number of neighbors, motion velocity and transmission environment of sensors. Meanwhile, a clustering function was constructed to optimize the clustering, eliminating extremely large or small clusters.Finally, the simulation proves that the DEAL protocol is more conducive to prolonging the life cycle of the sensor network. The SF-CHs algorithm can reduce the residual energy variance of nodes in the network, and the network failure time is later, which provides a way to improve the stability of the network and reduce energy loss.
Keywords: Wireless sensor network (WSN), ubiquitous power internet of things (UPIoT), cluster head (CH) selection, clustering optimization.

1 Introduction

The ubiquitous power Internet of Things is built around all aspects of the power system, fully utilizing modern information technologies such as mobile internet and artificial intelligence, and advanced communication technologies to realize the interconnection of all things and human-computer interaction in all aspects of the power system. It has comprehensive state perception, efficient information processing and application [18]. A smart service system with convenient and flexible features. It is a specific application of "5G" technology and the Internet of Things in the power industry system. It realizes a mode of interconnecting people-machine-devices inside and outside the power grid, and generates real-time shared data during the running process.Ubiquitous power Internet of Things (UPIoT) is a wireless sensor network (WSN) formed after applying the IoT technique in power systems. The wireless sensors in the network enable the information exchange between different links in power systems. The WSN is the key to the construction of the UPIoT, for it carries various strengths of distributed processing systems, including high monitoring accuracy, strong fault tolerance, large coverage, remote telemetry and control, self-organization and multi-hop routing [16]. In order to extend the network life cycle, the node energy in the sensor network can be balanced by randomly selecting the cluster head nodes to spread the energy consumption in the network.

However, there is a contradiction between the largescale and high complexity of the UPIoT and the limited energy of each sensor. When it transmits information to the sink, the sensor may lose efficacy due to excessive energy consumption, leading to the failure of the entire network. To extend the network lifecycle and balance sensor energy, the data transmitted by common sensors can be collected and integrated by a cluster head (CH), before being sent to the sink. This strategy eliminates the long-distance transmission between common sensors and the sink, and reduces the energy consumed in data transmission, thus extending the network lifecycle [5] [6].

Much research has been done at home and abroad on CH selection. For example, Reference [10] proposes a CH selection algorithm based on grid dynamic energy threshold, which selects CHs according to the sensor-grid center distance and the remaining energy of each sensor; considering network dynamics, the algorithm enhances the energy efficiency of sensors [19]. Reference [11] takes account of the energy consumed in data transmission of CH selection and information aggregation, and presents a reservation mechanism that lowers the energy consumption by controlling the number of messages generated in clustering and CH selection. Reference [13] designs a CH selection algorithm based on the fuzzy balance cost, which determines the how each sensor qualifies for CH by the sensor-sink distance and neighbor density, thus ensuring the load balance. Based on the dynamic characteristics of the nodes, residual energy and node location, the above papers provide an effective solution for the election of cluster heads, which plays an active role in improving the stability of wireless sensor networks and extending the life of the network. Overall, the existing studies on CH selection have the following defects:(1) The sensor's propagation rate and energy consumption will change in the harsh environment, and the life of the sensor will be affected. Therefore, the network environment factor should be considered in the process of selecting the cluster head; (2) In the current research, the ubiquitous power Internet of Things is not combined with the wireless sensor network, and the description of the ubiquitous power Internet of Things is also less. How to ensure the stability of the ubiquitous power Internet of Things operation needs to be solved urgently.

In the context of the UPIoT, this paper attempts to create a CH selection algorithm for the WSN to overcome the energy imbalance between sensors, which is common in WSNs [1] [2] [8] [9] [14]. Firstly, the author introduced the classic CH selection method called low energy adaptive clustering hierarchy (LEACH), and further optimized the selection results considering the remaining energy and sensor-sink distance, creating the distance and energy based advanced LEACH (DEAL) protocol [4] [17]. Next, the selectivity function-based CH selection (SF-CHs) algorithm was proposed based on sensor energy, the number of neighbors, sensor motion velocity, and transmission environment. To optimize the clustering results (clustering optimization), the clustering function was set up to prevent extremely large or small clusters. Finally, the simulation analysis shows that the DEAL protocol outperformed the LEACH protocol in extending network lifecycle and minimizing energy loss, and that the CH selected by the SF-CHs algorithm and the cluster allocated by the proposed function enhanced the stability and lifecycle of the WSN.

2 CH selection for the WSN

2.1 WSN

In the UPIoT, wireless sensors and other intelligent acquisition devices are applied to various links, ranging from grid infrastructure, energy storage stations, thermal power plants, photovoltaic power plants, wind farms to power users. These intelligent acquisition devices form a WSN, in which the links are called sensors [7] [15]. The wireless sensors are responsible for collecting information like power, device parameters, and environmental status [3]. During power dispatching, each sensor transmits its own information wirelessly to the control center at the

virtual power plant. Known as the sink, the virtual power plant is responsible for making a scheduling plan based on the load and power generation information, aiming to effectively improve the grid access of distributed energy and the proportion of the controllable capacity in the grid.

It is assumed that each WSN sensor has the following features:

(1) Each sensor is deployed randomly through the network, knowing nothing about its location.

(2) Each sensor has a limit on energy, which attenuates in data transmission and monitoring, and can sense its remaining energy.

(3) Each sensor is unique, work independently of others, and can communicate directly with the sink.

(4) Each sensor is free to send and receive data, and judge its distance to the data sender by the perceived signal strength.

(5) Located at a fixed position, the sink has an unlimited amount of energy and can communicate with all sensors.

The energy consumed to transmit information between sensors can be calculated by:

$$E_{Tx}(k,d) = \begin{cases} kE_{\text{elec}} + k\varepsilon_{fs}d^2, & d \le d_0 \\ kE_{\text{elec}} + k\varepsilon_{mp}d^4, & d > d_0 \end{cases} \tag{1}$$

$$E_{Rx}(k) = k \times E_{elec}$$

where E_{Tx} and E_{Rx} are the energies consumed by a sensor to send and receive information, respectively; k is the number of bits being transmitted; E_{elsc} is the energy needed to transmit 1 bit of information; ε_{fs} and ε_{mp} are the power amplification factors under the said two models, respectively; d is the sender-receiver distance; d_0 is the threshold (if $d \le d_0$, the free space model should be adopted, which assumes that the energy consumed to send information is positively correlated with d^2 ; if $d > d_0$, the multipath attenuation model should be adopted, which assumed that the energy consumed to send information is positively correlated with d^4), $d_0 = \sqrt{\frac{\varepsilon_{fs}}{\varepsilon_{mp}}}$.

In the UPIoT, the energy of a sensor far away from the virtual power plant may be attenuated to zero after multiple communication, owing to the limited amount of sensor energy. The attenuation could affect the lifecycle of the entire network. Thus, a clustering mechanism was introduced to the network, dividing the power network into several subsets. Each cluster consists of a CH and a few common sensors (Figure 1).

In a cluster, cluster members mainly perform information collection or information transmission, and communicate with cluster heads in a single-hop or multi-hop manner. The cluster head is the same as the ordinary node, but can function like a gateway. It has a certain data processing capability, and communicates with the base station through the network formed between the cluster head and the cluster head. In the data transmission process, the ordinary node sends its own information to the cluster head, and the cluster head fuses the data in the cluster and sends it to the aggregation node. For the monitoring sensor nodes of each device, the distance from the cluster head is shorter than that of the sink node, so it consumes less energy, thus prolonging the life of the network, and also ensuring that the data communication effectively covers the entire network, and is easier. Optimize management of nodes within the network.

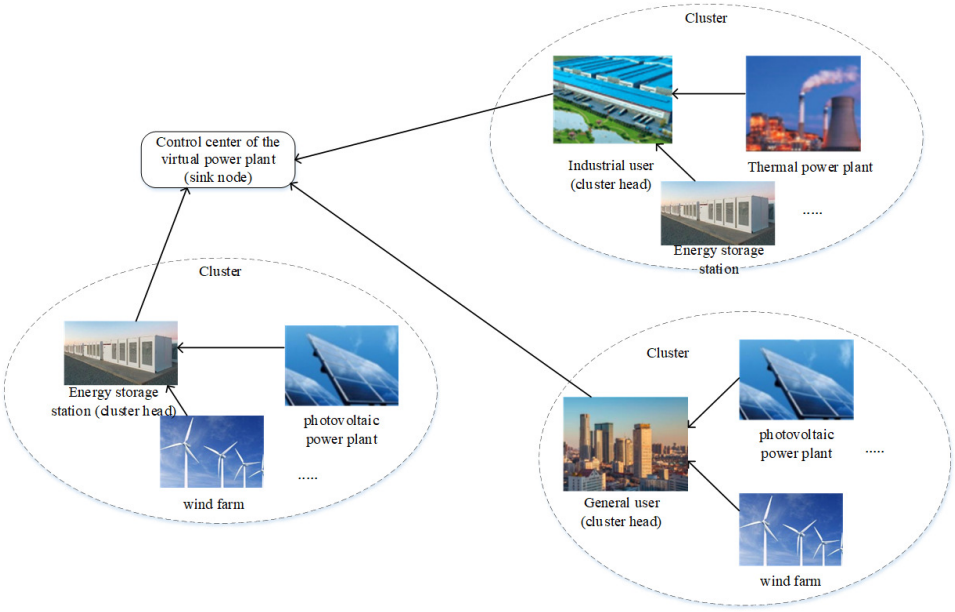


Figure 1: The WSN

2.2 CH selection

(1)DEAL

The LEACH is a typical clustering routing algorithm. Aiming to extend the lifecycle of the entire network, this protocol randomly selects CHs from the sensors that have never been CHs, such that the each sensor consumes a similar amount of energy.

In the LEACH protocol, the CHs selection is carried out in the following manner: each sensor randomly generates a number in the interval of [0, 1], and the sensors whose numbers are smaller than threshold T_i will be the CHs. The CH selection contains r_{total} cycles, where r_{total} is the ratio of the number of sensors divided by the number of CHs. During the selection process, every sensor has the potential to be the CH. The only difference lies in the number of cycles. Once selected, a sensor should not be CH again in the following cycles. The threshold T_i can be expressed as:

$$T(i) = \begin{cases} \frac{P}{1-P\times[r \bmod (1/P)]} & , i \in G \\ 0 & , \text{ Otherwise} \end{cases} \tag{2}$$

where P is the proportion of CHs in all sensors; mode is the remainder; r is the number of the current cycle; G is the set of sensors that have never been selected as CHs.

The LEACH algorithm effectively extends the network life cycle, but there are still some problems and some limitations:

1) Since LEACH assumes that all nodes can communicate directly with the aggregation node, and each node has the ability to support different MAC protocols, the protocol is not suitable for use in large-scale wireless sensor networks;

2) The protocol does not specify how the number of cluster head nodes can be distributed to the entire network. Therefore, it is very likely that the selected cluster head nodes are concentrated in a certain area of the network, so that there are no cluster head nodes around some nodes;

3) Since LEACH assumes that in the initial cluster head selection round, all nodes carry the same energy, and each node that becomes the cluster head consumes roughly the same amount

of energy. Therefore, the protocol is not suitable for networks with unbalanced node energy;

4) In the cluster head election, the impact of the death node on the network is not considered. The number of cluster heads does not change dynamically with the death of the network node. The number of cluster heads cannot always be optimal, and the death time will be earlier. Affect the stability of the network;

5) Uneven clustering can cause uneven distribution of network energy consumption. Single-hop transmission makes energy consumption of cluster head nodes far from the base station large.

As mentioned before, the DEAL is an advanced LEACH protocol through optimization based on distance and energy. It considers both the mean and variance of the remaining energy of each sensor. In the DEAL, the threshold for sensor i can be defined as:

$$T(i) = (1 - P) \times G_p + P \times CS_p \tag{3}$$

where P is the optimal proportion of CHs in all sensors; G_p and CS_p are the probabilities for a common sensor and the current sensor to be selected as CH, respectively. The G_p and CS_p can be calculated by formulas (4) and (5), respectively:

$$G_p = \frac{k}{N - k \times (rmod\frac{N}{k})} \tag{4}$$

where k is the number of CHs generated per cycle; N is the number of all network sensors; r is the number of the current cycle.

$$CS_p = \left(\frac{std\,(E_i)}{\frac{mean(E_i)}{Eo_i}} \right) - \left(P \times \frac{l_{n2B}}{l_{n2B}^{\text{farthest}}} \right) + \frac{N}{x_m \times y_m} \tag{5}$$

where $std(E_i)$ and $mean(E_i)$ are the mean and variance of the remaining energy of sensor i, respectively; Eo_i is the initial energy of sensor i; l_{i2B} is the distance from sensor i to the sink; $l_{n2B}^{farthest}$ is the distance from the sink to the farthest surviving sensor; $x_m \times y_m$ is the size of the entire network.

The DEAL retains the CH selection method of the LEACH: each sensor randomly generates a number in [0, 1], and the sensors whose numbers are smaller than the corresponding threshold will be the CHs. The optimization of the threshold function makes it more likely for sensors with high energy and close sensor-sink distance to be CHs. Thus, the CH selection is less random and the network can survive for a longer time.

(2) SF-CHs

In order to ensure the energy balance of sensor nodes, extend the life cycle of the network, and ensure the timely collection and effective management of power equipment information, this paper proposes a cluster head selection algorithm based on the selection function, and selects nodes with larger value of selection function in the network as the cluster head node, which effectively avoids the problem of node failure due to excessive energy consumption and prolongs the network life.

Definition 1. Selectivity function

The selectivity function can be established according to the remaining energy, the number of neighbors, the relative motion velocity, and transmission environment of each sensor:

$$maxZ = \lambda_1 \times E_{rest} + \lambda_2 \times n_i + \lambda_3 \times v_i + \lambda_4 \times \omega_r \tag{6}$$

where $\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 = 1$, and $\lambda_3 < 0$. If the current sensor has the largest value of the selectivity function, it should be elected as the CH.

The remaining energy of a sensor E_{rest} can be obtained as:

$$E_{\text{rest}} = \frac{E_{\text{current}}}{E_{\text{max}}} \tag{7}$$

where $E_{current}$ and E_{max} are the current remaining energy and the maximum initial energy of the sensor, respectively. The high energy sensor should be preferred in CH selection, because CH usually consumes more energy than common sensors. For example, the CH needs to collect the information about the power equipment from all common sensors in the cluster, sort out and integrate the information, and then send them to the sink.

Let n_i be the number of neighbors of a wireless sensor.

Definition 2. Neighbor

A neighbor refers to any sensor within the circle (radius: $R = \sqrt{\frac{S}{\pi \times Q}}$) centering on the current sensor. The expression of the radius can be proved as follows:
It is assumed that all sensors are located in a square with a side length of one, and all clusters have the same coverage.

Then, the optimal number of CHs in the network can be defined as:

$$Q = \sqrt{\frac{N \times d_0}{2\pi}} \frac{M}{d_{BS}^2} \tag{8}$$

where N is the number of network sensors; M is the side length of the network; d_{BS} is the mean sensor-sink distance. If the network remains stable, the number of CHs remains basically the same after the replacement in each cycle, and the communication radii of the CHs can cover the entire network. Suppose adjacent clusters do not overlap in their communication ranges between adjacent clusters, the network coverage S must satisfy:

$$Q \times \pi R^2 \geq S \tag{9}$$

Thus, we have:

$$R = \sqrt{\frac{S}{\pi \times Q}} \tag{10}$$

The sensors with many neighbors should be preferred in CH selection, because the distance to CH is positively correlated with the energy loss in information transmission.

Let v_i be the motion velocity of the sensor i. In the WSN, some sensors are constantly moving. If these sensors be CHs, the stability of the network may be dampened. What is worse, such a CH may enter the adjacent clusters, increasing its distance to the sensors in its cluster and to the sink. In this case, more energy will be consumed to transmit the same amount of data, and the data may even get lost. Considering the negative effects of the moving sensors, the author set up a weighting factor $\lambda_3 < 0$ for the motion velocity of the sensors. The sensors with a low motion velocity will be preferred in the selection process.

Letω_r be the environmental impact factor of the transmission process. This factor was introduced to the selectivity function to describe how the natural conditions affect the wireless communication bandwidth and thus the data transmission rate, as well as the impacts of other unknown, uncontrollable changes.

3 Clustering optimization

3.1 Extremely large and small clusters

Let $X = \{X_1, X_2, X_3 \cdots X_n\}$ be the CHs being selected, with n be the number of clusters. The remaining common nodes determine the joined cluster according to the clustering function, and communicate with the cluster head of the cluster. After receiving the "Request to Join" message of all member nodes, the cluster head allocates the order in which the member nodes in the cluster send data, calculates the sending time of each cluster member, and sends the generated TDMA message to all member nodes in the cluster. The member node receives the message and saves its own time slot, and it is the turn to communicate with the cluster head when it sends itself.

Definition 3. Clustering function

Each sensor selects its cluster according to its distance to and the energy of each CH. In general, a sensor prefers a nearby CH with the most remaining energy. The clustering function can be expressed as:

$$minR_i = \alpha \times d_{ij} - \beta \times E_{rest}^{x_j} \tag{11}$$

where α and β are the influence factors of distance and energy, respectively; d_{ij} is the distance between sensor i and CH j; $E_{rest}^{x_j}$ is the remaining energy of CH j. Sensor i will be allocated to the cluster of the CH that minimizes the value of the clustering function.

It is difficult to predict the size and number of sensors of each cluster, as the CHs generated in each cycle differ in remaining energy and have random positions. This may lead to extremely large and small clusters (Figure 2).

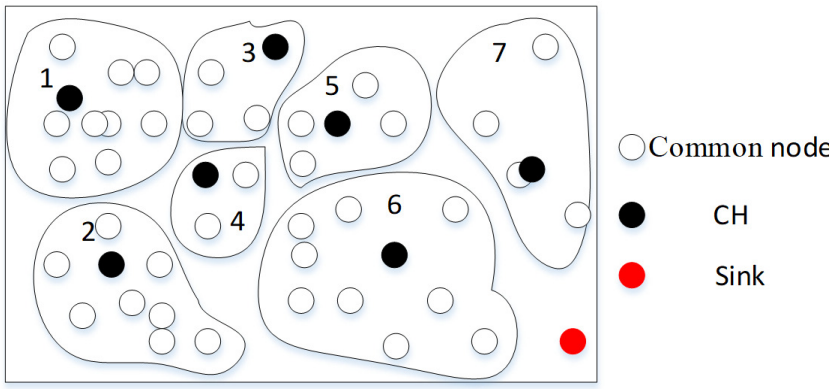


Figure 2: Extremely large and small clusters

As shown in Figure 2, the randomly distributed sensors can be divided into 7 clusters, in which clusters 3, 4, 5 and 7 are extremely small clusters while clusters 1, 2 and 6 are extremely large clusters. For each extremely large cluster, more energy is consumed as each sensor transmits its data to the CH, and as the CH integrates the received information; for each extremely small cluster, the CH is far away from the sink, pushing up the energy consumption in data transmission to the sink; Also, because the number of nodes in the cluster is small, the time for collecting and processing data is much shorter than that of extremely large cluster, so the data is sent to the sink node more frequently, resulting in excessive consumption of extremely small clusters and unbalanced energy consumption; in addition, extremely small clusters frequently send data to the sink node, and the sink node needs to participate in the collaborative acceptance of data, thereby

occupying the sink node more frequently, so that other cluster head nodes cannot cooperate with the sink node when data needs to be transmitted, thereby a conflict of collaboration has arisen. Therefore, both extremely large and small clusters may cause energy depletion and premature death of the CHs, resulting in a shorter network lifecycle. To solve the problem, the maximum and minimum capacity constraints can be defined for each cluster:

$$CA_j \geq N_{av} + \eta(E_{rest}^{x_j} - E_{rest}^{av}) \times \frac{N_{av}}{E_{rest}^{av}} \tag{12}$$

$$CA_j \leq N_{av} + \eta(E_{rest}^{x_j} - E_{rest}^{av}) \times N_{av} \tag{13}$$

where CA_j is the number of sensors in the cluster of CHj;N_{av} is the mean number of sensors of all clusters in the network; $E_{rest}^{x_j}$ is the remaining energy of CHj; E_{rest}^{av} is the mean remaining energy of all CHs; $0 < \eta < 1$ is the relationship factor between remaining energy and the number of sensors.

3.2 Data transmission procedure

The data transmission process in the network was divided into cycles. Each cycle was further splits into the cluster formation stage and the stable transmission stage. The former involves CH selection and clustering. The specific process is shown in Figure 3.

(1) Cluster formation stage. Before the first cycle, the sink confirms all surviving network sensors, while each sensor transmits its information about the initial energy to the sink, and the CHs are determined according to the best number of CHs and the selectivity function. The sensors elected as CHs will broadcast its information to the entire network, so that each common sensor selects its desired cluster according to the clustering function. The common sensor will send the clustering request to the corresponding CH, applying to join the cluster. The CH will always keep its receiver open. If the cluster being applied is saturated with sensors, the common sensor will apply to enter the cluster with the next optimal conditions. If failing to join any cluster after the end of the current cycle, the general cluster will form a cluster itself. After confirming the sensors accepted to its cluster, the CH will return an acceptance signal to these sensors and assign a time interval for communication to each of these sensors. In addition, the CH will also create a time division multiple access (TDMA) schedule, and broadcast it across the network. After that, the common sensors will send data according to the schedule, preventing timing disorder in the transmission process.

In order to reduce energy consumption, each cluster node uses the energy control mechanism to set the amount of energy used for transmission according to the strength of the cluster head transmission signal. When the time slot is not in turn, the nodes in the cluster can turn off the wireless communication module to save energy, and the cluster head node must remain active to receive data.

(2) Stable transmission stage. In the stable transmission stage, each common sensor transmits its monitored information to the corresponding CH. After receiving all the information, the CH will integrate the information and send the result to the sink. The data fusing process can remove unnecessary information, reduce traffic and energy consumption, and extend the network cycle. During data transmission, the time is divided into a number of intervals according to the TDMA schedule, so that each intra-cluster sensor can send data to the CH at its time interval. If a part of data is not sent in the current time interval, it should be transmitted gain in the next time interval.

After receiving certain data, the cluster head node integrates the data, strengthens meaningful signals, reduces uncorrelated noise, and sends the data to the sink node. This mode can

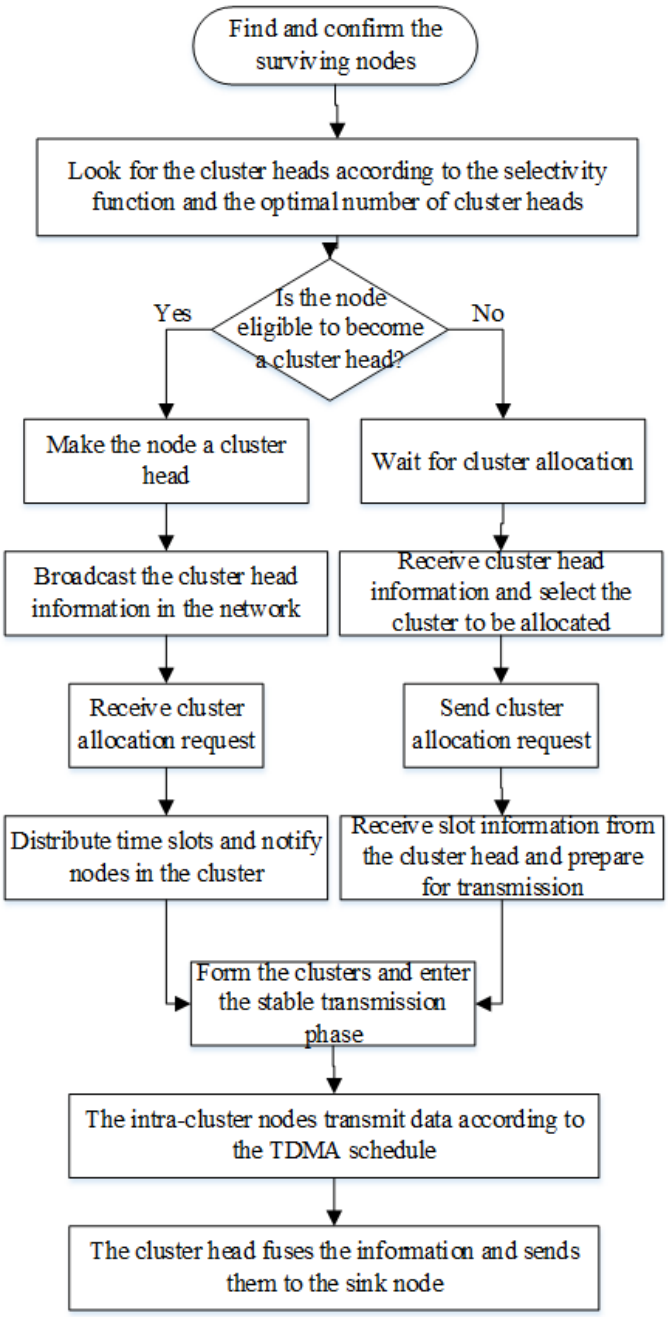


Figure 3: Data transmission procedure

reasonably reduce the traffic, but when the cluster head is far away from the sink node and the data packet is large, the energy of the cluster head node is quickly consumed.

In order to reduce the inter-cluster interference, the DSss mechanism is used for inter-cluster communication. The intra-cluster nodes use a unique spreading code to transmit data to the cluster head. The cluster head uses the spreading code to filter all received energy. In order to reduce inter-cluster interference and its own energy consumption, the node can adjust the transmit power by itself. The cluster head sends data to the sink node using a fixed spreading code. The cluster head transmits the data using the CSMA sounding channel mechanism. If another cluster is transmitting data using the spreading code, the cluster head waits for other

cluster heads to complete the data transmission; if not, the cluster head node directly uses the spread spectrum based The code is used for data transmission. After the data is stably transmitted for a period of time, the network restarts a round of cluster head elections, and the above process is repeated.

4 Simulation experiment

4.1 Parameter setting

The three CH selection algorithms, namely, the LEACH, the DEAL and the SF-CHs, were simulated on the Matlab, aiming to test the algorithm's network lifetime and network performance. The network coverage was set to $100m \times 100m$, with the sink at the center. All the sensors were evenly distributed in the network, all of which know the coordinates of the sink. The initial energy of each sensor was set to 1J, the data packet size was set to 4,000 bits, the energy consumed to transmit 1 bit of data E_{elec} was set to 50nJ/bit, the ε_{fs} and ε_{mp} were set to $10pJ/bit/m^2$ and $0.0013pJ/bit/m^4$, respectively, and the number of cycles were set to 300. After each cycle, the energy of each sensor will be attenuated to a certain extent. If the energy of a sensor falls to zero, the sensor can be deemed as failed. Multiple sensor failures will make the monitoring ineffective. Hence, the end of the network lifecycle was defined as the failure of 95% of all network sensors [10].

4.2 Results analysis

The three algorithms were simulated under a 200-sensor network and a 300-sensor network. The CH selection and clustering results are shown in Figure 4 below.

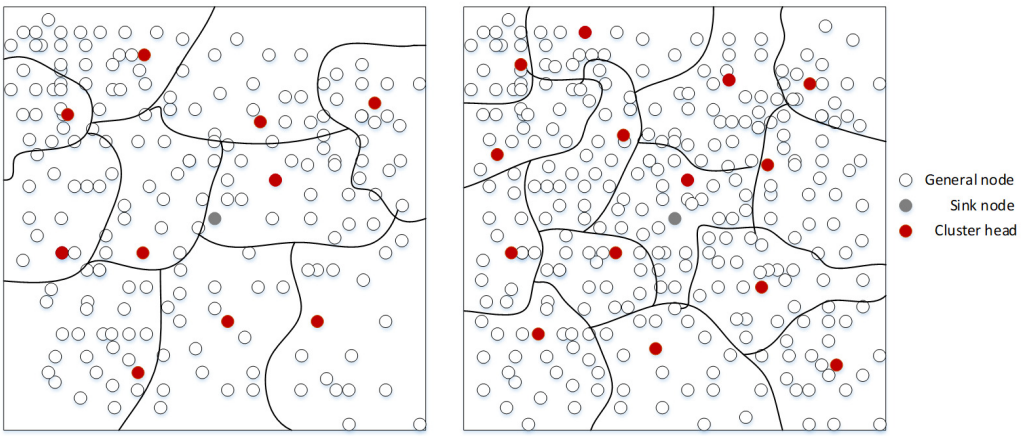


Figure 4: Clustering results

It can be seen from that the sink was located at the network center, and the sensors were clustered relatively evenly, the number of nodes in each cluster is within the range of suitable cluster capacity, with no extremely large or small clusters. The clusters had no significant difference in energy loss.

Taking the 200-sensor network as an example, the relationships between the number of failed sensors and the number of cycles of the three algorithms are presented in Figure 5. Under the operation conditions of the LEACH algorithm, the first sensor failure appeared in cycle 25, and the network lifecycle ended at cycle 225. Under the operation conditions of the DEAL, the first sensor failure occurred in cycle 35, and the network stopped operation in cycle 231. Under the

CH selection and clustering by the SF-CHs, the first sensor failure was observed in cycle 42 and the network lifecycle terminated at cycle 275.

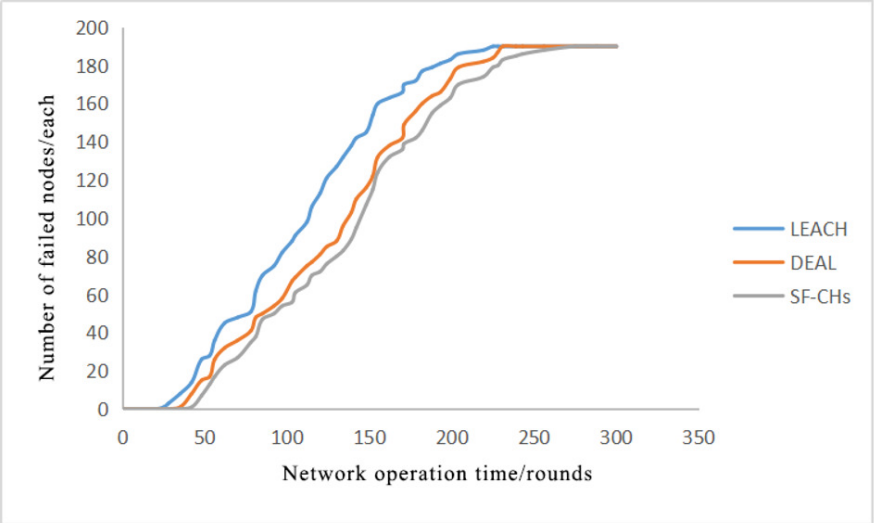


Figure 5: Number of failed sensors of the three algorithms

The above results show that the network lifecycle ended at basically the same moment under the DEAL and the LEACH. However, the first sensor failure appeared later and the network survived longer under the DEAL. Under the proposed SF-CHs, the network lifecycle was longer than that under the LEACH and the DEAL, the death of the first sensor was further delayed, and the network lifecycle was extended by 16.5%. To sum up, the proposed SF-CHs can ensure the balance of energy consumption and extend network lifecycle, thanks to the consideration of neighbors, motion velocity and transmission medium and the optimization of the clustering process.

In order to study the stability of the network, this paper compares the average residual energy of the nodes in the network and the variance of the residual energy of the nodes under the three algorithms. The results are shown in Figure 6 and Figure 7.

Figure 6 compares the residual energy of the network under three different algorithms, reflecting the change in energy consumption rate. When using the LEACH algorithm, due to poor cluster head selection, the energy consumption of the previous network is large, and the slope of the curve is large. In the later stage, the number of nodes in the network is reduced, the network energy consumption is also reduced, and the slope is reduced. The DEAL algorithm is used for the LEACH algorithm. Optimized, so the energy consumption is lower, the slope of the curve is smaller and smaller as the surviving node is reduced. The SF-CHs algorithm is not only better in cluster head selection, but also better clustering result. The consumption is basically the same, the curve tends to be a straight line, and the same, the overall energy consumption is reduced due to node death in the later stage.

As shown in Figure 7, the LEACH brought about the maximum variance of remaining energy, while the SF-CHs minimized the variance, and the variance is reduced by 32%. This means the DEAL protocol reduces the variance of remaining energy from the level of the LEACH, while the proposed SF-CHs can effectively enhance network stability and extend network lifecycle.

Figure 8 compares the energy values consumed per round under the operating conditions of the three algorithms. The energy consumption per round reflects the cluster head selection and clustering results. When the cluster head selection is correct, the energy consumed per

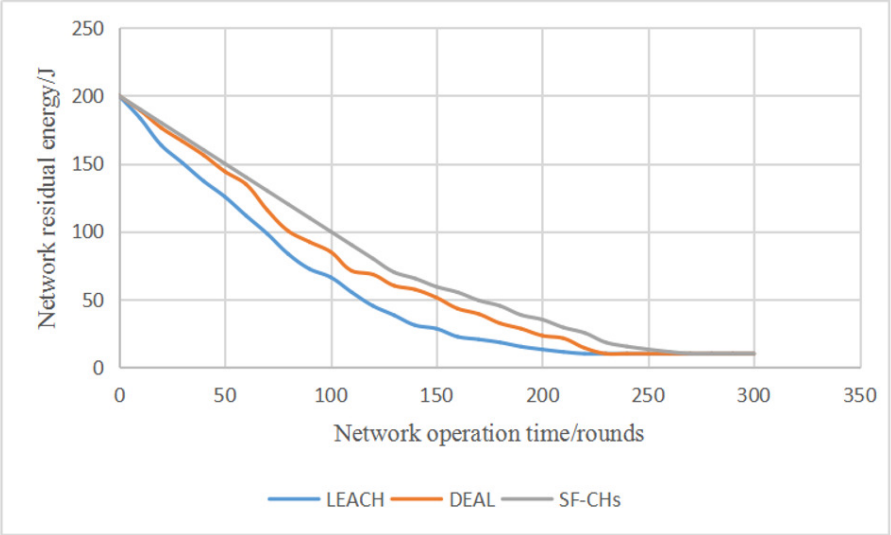


Figure 6: Network residual energy comparison curve

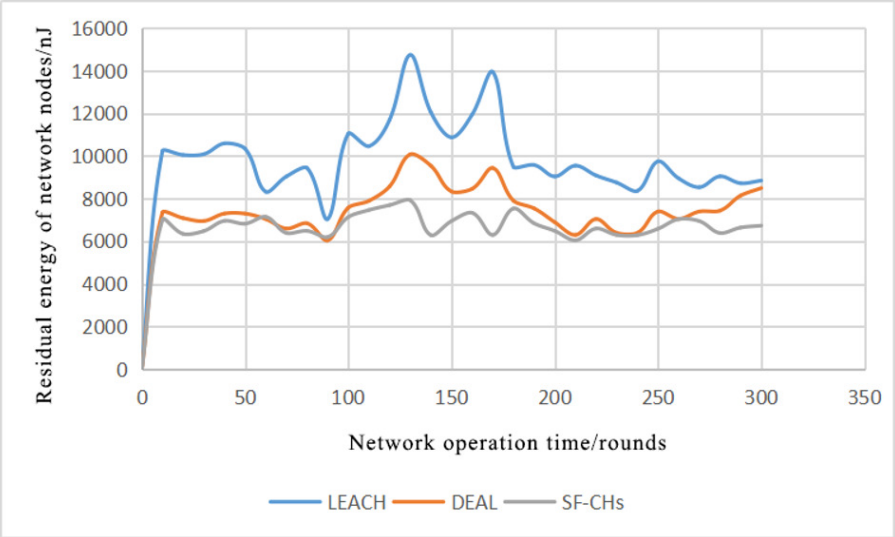


Figure 7: The variance of remaining energy of sensors of the three algorithms

round is naturally less. Large networks typically consume less energy than smaller networks. As shown in Figure 7, based on the optimized LEACH algorithm, the DEAL algorithm consumes less energy than the LEACH algorithm. Therefore, it can be seen that the DEAL algorithm has a better selection of cluster heads. In addition, before 240 rounds, the SF-CHs algorithm The energy consumption per round is much less, indicating that the clustering result of the SF-CHs algorithm is better; at the end of the running cycle, the energy consumption per round of SF-CHs and DEAL is higher than that of the LEACH algorithm. The reason may be In the LEACH algorithm, the node is too dead, so it consumes less energy, while in the other two algorithms, there are more surviving nodes, which naturally consume more energy.

In addition, this paper analyzes how the location of the aggregation node affects the number of dead nodes, as shown in Figure 9.

The two cases are simulated separately. One case is that the aggregation node is located at the center of the network, and the other case is that the aggregation node is located at the edge

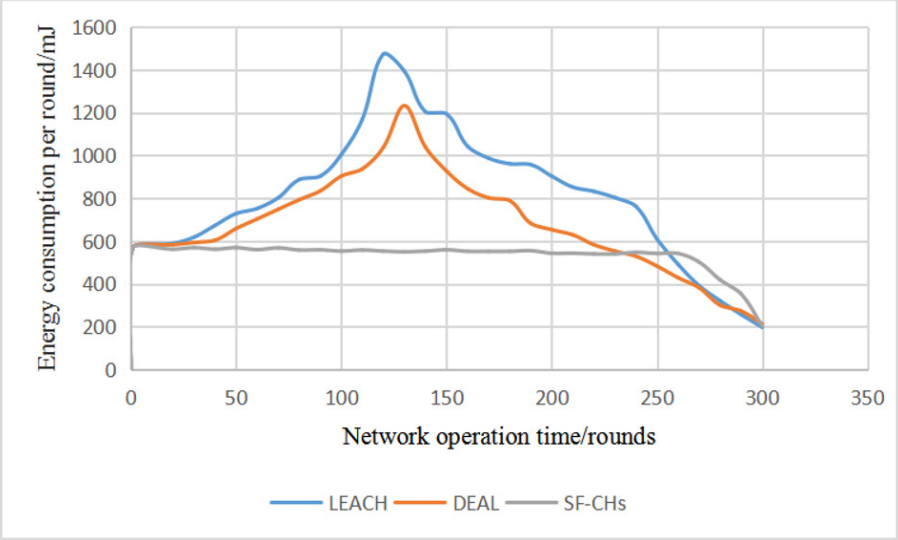


Figure 8: Comparison of energy consumption per round under the three algorithms

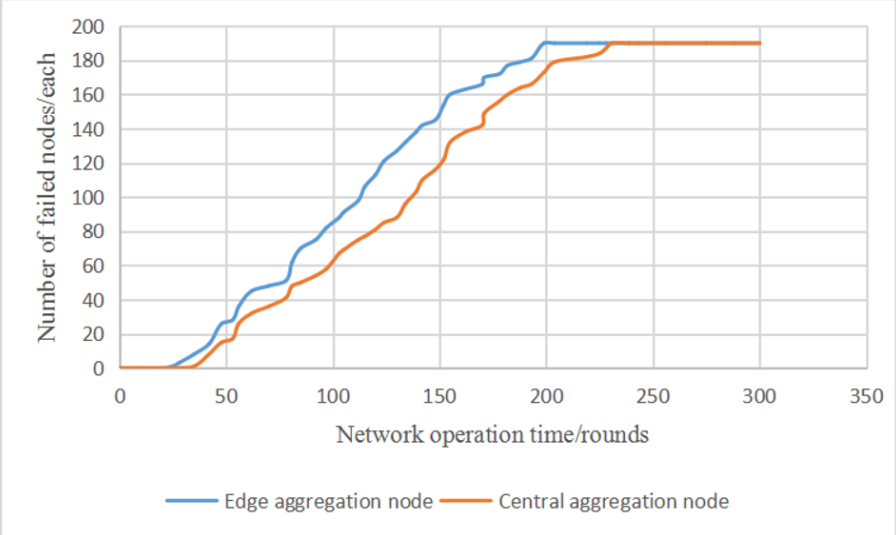


Figure 9: Number of dead nodes when the sink node is in different locations

of the network. It can be seen from the figure that the aggregation node has a dead node in the network center and the network failure time is later, and the network life is longer.

Based on the above analysis, the proposed SF-CHs algorithm is beneficial to improve network stability and extend network lifetime, and can play a positive role in the development of ubiquitous power Internet of Things.

5 Conclusion

The UPIoT is an important research topic of the future power industry. The performance of the UPIoT depends heavily on wireless sensors. Improving the stability and survival period of wireless sensor networks can better promote the development of ubiquitous power Internet of Things. To solve the energy depletion of wireless sensors in data transmission, a common

method is to set up CHs and divide WSN sensors into several clusters, thus enhancing the energy efficiency. In this paper, a WSN is set up in the context of the UPIoT, taking the application links of the power system as sensors, and the information exchange between the sensors is realized through sensors. On this basis, the basic principle of the LEACH and DEAL protocols was explained: each sensor generates a random number, and the sensors whose numbers are smaller than the corresponding threshold will be the CHs. Next, the SF-CHs algorithm was designed in two steps: the cluster selectivity function was created based on the remaining energy, number of neighbors, motion velocity and data transmission environment of each sensor, and the clustering function was established to optimize the clustering method, eliminating extremely large or small clusters. Finally, the simulation shows that compared with the LEACH protocol and the DEAL protocol, the SF-CHs algorithm has a late failure node, the network lifetime is extended by 16.5%, and the variance of the remaining energy of the nodes in the network is reduced by 32%, which effectively improves the stability of the network. In addition, when the aggregation node is located in the center of the network, the time of the failed node is later, and the network life is longer. Therefore, the SF-CHs algorithm can provide a solution for extending the network life and improving the stability of the network.

The research work of the thesis is based on the ubiquitous power internet of things, but all the research is only based on the theoretical level. The simulation results are only verified in the virtual environment. The future research work will apply the research results to the actual environment for verification. Development and application verification for specific wireless sensor network monitoring environment.

Funding

This work is supported by the Humanities and Social Sciences Project of Chinese Ministry of Education (Grant No.: 17YJCZH062).

Author contributions. Conflict of interest

The authors contributed equally to this work. The authors declare no conflict of interest.

Bibliography

[1] Aghera, K.; Pambhar, H.; Tada, N. (2017). MMR-LEACH: Multitier multi-hop routing in LEACH protocol, *Proceedings of International Conference on Communication and Networks*, 205–214, 2017.

[2] Amirthalingam, K.; Anuratha, V. (2017). Improved LEACH: A modified LEACH for wireless sensor network, *IEEE International Conference on Advances in Computer Applications*, 255–258, 2017.

[3] Awad, F. H. (2018). Optimization of relay node deployment for multisource multipath routing in wireless multimedia sensor networks using gaussian distribution, *Computer Networks*, 145, 96-106, 2018.

[4] Bao, X.; Xie, J.; Nan, L.; Li, S. (2014). WRECS: An improved cluster heads selection algorithm for WSNs, *Journal of Software*, 9(2), 78-89, 2014.

[5] Baranidharan, B.; Santhi, B. (2016). DUCF: Distributed load balancing unequal clustering in wireless sensor networks using fuzzy approach, *Applied Soft Computing*, 40, 495-506, 2016.

[6] Batra, P. K.; Kant, K. (2016). LEACH-MAC: A new cluster head selection algorithm for Wireless Sensor Networks, *Wireless Networks*, 22(1), 49-60, 2016.

[7] Chidean, M. I.; Morgado, E.; Del Arco, E.; Ramiro-Bargueno, J., Caamano, A. (2015). Scalable data-coupled clustering for large scale WSN, *IEEE Transactions on Wireless Communications*, 14(9), 4681-4694, 2015.

[8] Jia, Y. L.; Chang, X. M. (2017). Cluster heads selection algorithm for wireless sensor networks based on cluster heads sending energy consumption, *Computer Engineering and Applications*, 53(22), 82-86, 2017.

[9] Lee, J. Y.; Jung, K. D.; Moon, S. J.; Jeong, H. Y. (2017). Improvement on LEACH protocol of a wide-area wireless sensor network, *Multimedia Tools & Application*, 76(19), 19843-19860, 2017.

[10] Liang, P.; He, W. (2017). Grid dynamic energy threshold-based cluster header algorithm in wireless sensor network, *Chinese Journal of Sensors and Actuators*, 30(10), 1583-1588, 2017.

[11] Mehra, P. S.; Doja, M. N.; Alam, B. (2018). Fuzzy based enhanced cluster head selection (FBECS) for WSN, *Journal of King Saud University - Science*, 2018.

[12] Mittal, N. (2019). Moth flame optimization based energy efficient stable clustered routing approach for wireless sensor networks, *Wireless Personal Communications*, 104(2), 677-694, 2019.

[13] Shalabi, M.; Anbar, M.; Wan, T. C.; Khasawneh, A. (2018). Variants of the low-energy adaptive clustering hierarchy protocol: Survey, issues and challenges, *Electronics*, 7(8), 136-164, 2018.

[14] Snigdh, I.; Gupta, N. (2016). Quality of service metrics in wireless sensor networks: A survey, *Journal of the Institrtion of Engineers*, 97(1), 91-96, 2016.

[15] Srikanth, B.; Kumar, H.; Rao, K. U. M. (2018). A robust approach for WSN localization for underground coal mine monitoring using improved RSSI technique, *Mathematical Modelling of Engineering Problems*, 5(3), 225-231, 2018.

[16] Sun, L. M.; Li, J. Z.; Chen, Y. (2005). Wireless sensor networks, *Beijing: Tsinghua university press*, 1-5, 2005.

[17] Thakkar, A. (2017). DEAL: Distance and energy based advanced LEACH protocol, *International Conference on Information and Communication Technology for Intelligent Systems*, 370-376, 2017.

[18] Wei, X. (2014). Power wireless sensor network clustering routing optimization algorithm research, *North China electric power university (Beijing)*, 2014.

[19] Zahedi, A.; Arghavani, M.; Parandin, F.; Arghavani, A. (2018). Energy efficient reservation-based cluster head selection in WSNs, *Wireless Personal Communications Wireless Personal Communications*, 100(3), 667-679, 2018.

INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL
ISSN 1841-9836, e-ISSN 1841-9844, 14(3), 359-374, June 2019.

Effect of Soft Errors in Iterative Learning Control and Compensation using Cross-layer Approach

G.-M. Jeong, K. Lee, S.-I. Choi, S.-H. Ji, N. Dutt

Gu-Min Jeong
School of Electrical Engineering
Kookmin University, Seoul, Korea
gm1004@kookmin.ac.kr

Kyoungwoo Lee*
Department of Computer Science
Yonsei University, Seoul, Korea
*Corresponding author: kyoungwoo.lee@gmail.com

Sang-Il Choi
Department Applied Computer Engineering
Dankook University, Korea
choisi@dankook.ac.kr

Sang-Hoon Ji
Advanced Robotics Group
KITECH, Ansan, Korea
robot91@kitehc.re.kr

Nikil Dutt
Department of Computer Science
UC Irvine, CA, USA
dutt@ics.uci.edu

Abstract: In this paper, we study the effects of radiation-induced soft errors in iterative learning control(ILC) and present the compensation techniques to make the ILC systems robust against soft errors. Soft errors are transient faults, which occur temporarily in memories where the energetic particles strike the sensitive region in the transistors mainly under abnormal conditions such as high radiation, high temperature, and high pressure. These soft errors can cause bit value changes without any notification to the controller, affect the stability of the system, and result in catastrophic consequences. First, we investigate and analyze the effects of soft errors in the ILC systems. Our analytical study shows that when a single soft error occurs in the output data from the ILC, the performance of the learning control is significantly degraded. Second, we propose novel learning methods by incorporating the existing techniques across the system abstraction levels in the ILC to compensate for soft-error-induced incorrect output. The occurrence of soft errors is estimated by using a monotonic convergence of the erroneous outputs in a cross-layer manner, and our proposed methods can significantly reduce these negative impacts on the system performance. Under the assumption of soft error occurrence, our analytic study has proved the convergence of the proposed methods in the ILC systems and our simulation results show the effectiveness of the proposed methods to efficiently reduce the impacts of soft-error-induced outputs in the ILC systems.
Keywords: Soft error, iterative learning control, compensation.

1 Introduction

With the advances in the highly integrated semiconductor and microelectronic manufacturing techniques, the variability of electronic devices has drastically increased in real time. On the

other hand, soft errors or transient faults induced by various sources (e.g., heat and cosmic radiation) have become one of the critical threats to the system stability [4]. Soft errors, different from hard errors, which are permanent faults in the system and cannot be corrected without replacement, are difficult to detect and can be as critical as hard errors to the system reliability in certain application areas [4, 10, 14, 15, 17, 19, 22]. The probability of soft error occurrence can be extremely low because these errors generally occur under abnormal situations. However, a single soft error can result in a bit value change, that is, from 0 to 1, or vice versa, in the memory without notifications to the application if there is no error detection or correction mechanism. Thus, depending on the bit location of the soft error, it can be critical to the stability of the controller.

Until now, there have been several researches for fault tolerant techniques to reduce the negative impacts of the soft errors on embedded systems. In particular, cross-layer-based approaches have been considered as promising techniques to analyze the trade-off among the performance, power, and reliability and have been presented to maximize the reliability with minimal overheads of the power and performance in various embedded systems [3, 8, 12, 13, 21].

Iterative learning control(ILC) systems have been widely deployed and used for various conventional and emerging control applications [1, 2, 5, 7, 9, 16, 20]. A single error or an incorrect output can cause a catastrophic consequence because these control systems can be used for mission- and safety-critical systems. However, no research has been conducted to protect the ILC systems from soft errors, to the best of our knowledge, while the soft error rate is exponentially increasing with technology scaling. To this end, we analyze the effects of soft errors in the ILC and propose compensation methods to increase the reliability of the ILC systems against soft errors, in particular, techniques in a cross-layer manner. First, we discuss the impacts of a single soft error on the output in the ILC systems and show that the soft error significantly affects the performance of the learning control. For an intensive analysis, we have selected an example with 4-byte floating point number [6] and have shown that a soft error in the ILC systems can lead to a large number of errors in the value and incorrect outputs eventually as threats to the system stability and reliability. Indeed, a soft error can significantly reduce the speed of the learning curve in a linear discrete-time system. Therefore, it is required to design and develop mitigation techniques to protect the ILC systems from soft errors. Second, we propose novel compensation methods against soft errors for the ILC systems. We have made a few assumptions about the soft errors in the ILC systems and investigated the new detection and compensation methods. In particular, from a monotonic convergence of the output error, we have estimated the occurrence of the critical soft errors that significantly affect the system output and result in catastrophic consequences. We propose four compensation methods: (i) exact rollback method, (ii) recovery with the previous iteration value, (iii) recovery with the adjacent value, and (iv) recovery with the desired value. In addition, the methodology and effectiveness of these four compensation techniques have been evaluated on the basis of their advantages and disadvantages. Based on the assumptions that we have made for the soft errors in this study, the convergence of each method has been presented. Our extensive simulation results show that the negative impacts of the soft errors can be successfully canceled with the proposed compensation methods.

The remainder of this paper is organized as follows. In Section 2, related works are briefly summarized for the soft-error related problems and the ILC for discrete-time systems. In Section 3 and 4, the effect of the soft error in the ILC is discussed and a new learning method to compensate for the effect of the soft error is introduced, respectively. An illustrative example is described in Section 5 and the conclusion follows in Section 6.

2 Related works

In this section, the preliminary results for the researches on the soft-errors and the ILC are briefly discussed.

2.1 Soft errors

Several decades of technology scaling have brought us to a point where transistors have become extremely susceptible to even small fluctuations in the voltage levels, slight noise in the power supply, signal interference, and cosmic particle strikes [4]. Any of these effects can temporarily toggle the logic value of a transistor, and therefore, it is called a transient fault or soft error. These soft errors are not permanent and nondestructive, that is, resetting of the device restores the normal behavior. An energetic particle such as an alpha particle, a neutron, or a free proton, can form a diffusion region of a CMOS transistor and produce a charge that can result in toggling of the logic value of the gates or flip-flops. This phenomenon of change in the logic state is called a transient fault or a soft error. Soft errors can result in erroneous program states, incorrect outputs, and, eventually, system crashes.

Soft errors have already been proved to cause significant fiscal damages [22]. For example, SUN blamed the soft errors for the crash of their million-dollar line SUN flagship servers [14]. More recently, Hewlett Packard acknowledged that a large installed base of a 1024-CPU server system in the Los Alamos National Laboratory has been frequently crashing owing to soft errors [15]. In another incident, the soft errors brought a billion-dollar automotive factory to halt every month [10], [17]. Further, highly integrated reliability-sensitive embedded devices such as mobile health-care systems and anti-lock braking systems (ABSs) in the automotive engine control units (ECUs) are significantly threatened by the exponentially increasing soft error rates with technology scaling. Thus, it is necessary to combat the soft errors in embedded systems in both emerging and traditional computing environments.

Recently, several selective protection techniques have been proposed in order to combat soft errors in a cost-efficient manner by protecting the failure critical data only [11, 13, 18, 21]. For instance, partially protected caches(PPC) [11] utilizes the knowledge of the content and device hardware capabilities to selectively store the critical data in a more reliable hardware (e.g., protected cache) and non-critical data in a less reliable one (e.g., non-protected cache). In [3], a novel memory cell design method considering robustness to soft-errors has been proposed for a deep learning accelerator.

Cross-layer-based techniques have been presented to maximize protection at a minimal cost by coupling and cooperating different error control schemes across the system abstraction layers from the hardware layer to the application layer in mobile embedded systems [12]. In [8], a cross-layer approach has been introduced to robust face recognition against soft errors. However, these techniques have been proposed to mitigate the impacts of the soft errors on domain specific embedded systems or general computer systems, and not control systems, in particular.

2.2 ILC for linear discrete time system

Some of the preliminary results for the ILC are briefly discussed, considering the monotonic convergence of the output errors [5]- [9].

We consider a linear time invariant(LTI) system described by

$$\begin{aligned} x(i+1) &= Ax(i) + Bu(i) \\ y(i) &= Cx(i) \end{aligned} \tag{1}$$

where, $u \in \mathbb{R}^1$, $x = [x_1, \cdots, x_n]^T \in \mathbb{R}^n$, and $y \in \mathbb{R}^1$ are the input, the state, and the output of the system, respectively. A, B and C are matrices of appropriate dimensions. Let $u^d(i)$, $x^d(i)$, and $y^d(i)$ represent the input, the state, and the output corresponding to the desired trajectory. Also let the desired output $y^d(i), i \in [\sigma, N+\sigma-1]$, be given and $\mathbf{u}_{[i,j]} := [u(i), \cdots, u(j)]^T, \mathbf{y}_{[i,j]} := [y(i), \cdots, y(j)]^T$. Here σ denotes the relative degree, which means that $CA^{\sigma-1}B \neq 0$ and $CA^iB = 0, i < \sigma - 1$.

The transfer function of the system is represented as

$$G(z) = \frac{\beta_1 z^{n-1} + \cdots + \beta_n}{z^n + \alpha_1 z^{n-1} + \cdots + \alpha_n}. \tag{2}$$

Here, we assume that the relative degree, σ, is known as *a priori* (i.e., $\beta_1 = \cdots = \beta_{\sigma-1} = 0$).

Based on the relative degree, we can obtain the input-to-output relation as

$$y(i+\sigma) = CA^\sigma x(i) + CA^{\sigma-1}Bu(i). \tag{3}$$

In addition, the following input-to-output mapping is used for ILC:

$$\begin{aligned} \mathbf{y}_{[\sigma,N+\sigma-1]} &= \mathbf{H}x(0) + \mathbf{J}\mathbf{u}_{[0,N-1]}, \\ \mathbf{H} &= \left[\ (H_0)^T \ , \ \cdots \ , \ (H_{N-1})^T \ \right]^T, \\ \mathbf{J}_{d_0} &= \begin{bmatrix} J_0 & 0 & \cdots & 0 \\ J_1 & J_0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ J_{N-1} & J_{N-2} & \cdots & J_0 \end{bmatrix}, \end{aligned} \tag{4}$$

where $H_l = CA^\sigma$ and $J_l = CA^{\sigma+l-1}B$. The time interval for the output of interest is $[\sigma, N+\sigma-1]$ for minimum phase systems. Also, at every iteration, we set $x^k(0) = x^d(0)$.

Considering (4), an input update law is given as

$$\mathbf{u}^{k+1}_{[0,N-1]} = \mathbf{u}^k_{[0,N-1]} + \mathbf{S}\mathbf{e}^k_{[\sigma,N+\sigma-1]}, \tag{5}$$

where $\mathbf{e}^k_{[l,m]} = \mathbf{y}^d_{[l,m]} - \mathbf{y}^k_{[l,m]}$ and $\mathbf{S} \in \mathbb{R}^{N \times N}$ is the learning gain matrix.

The monotonic convergence of the above ILC method is shown in the following lemma.

Lemma 1. *For the uncertain system (1), if the condition*

$$||I - \mathbf{J}\mathbf{S}|| \leq \rho < 1, \tag{6}$$

holds for all k, the output error $\mathbf{e}^k_{[\sigma,N+\sigma-1]}$ monotonically converges to 0 as $k \to \infty$. That is,

$$||\mathbf{e}^{k+1}_{[\sigma,N+\sigma-1]}|| \leq \rho ||\mathbf{e}^k_{[\sigma,N+\sigma-1]}||, \lim_{k\to\infty} ||\mathbf{e}^k_{[\sigma,N+\sigma-1]}|| = 0. \tag{7}$$

Proof: From (4) and (5), we can obtain

$$\begin{aligned} \mathbf{e}^{k+1}_{[\sigma,N+\sigma-1]} &= \mathbf{J}\mathbf{u}^d_{[0,N-1]} - \mathbf{J}\mathbf{u}^{k+1}_{[0,N-1]} \\ &= \mathbf{J}\mathbf{u}^d_{[0,N-1]} - \mathbf{J}(\mathbf{u}^k_{[0,N-1]} + \mathbf{S}\mathbf{e}^k_{[\sigma,N+\sigma-1]}) \\ &= [I - \mathbf{J}\mathbf{S}]\mathbf{e}^k_{[\sigma,N+\sigma-1]}. \end{aligned}$$

Taking the norms on the both sides and from (6), we can obtain (7). □

3 Effect of soft error in iterative learning control

In this section, we analyze the effect of soft errors on the ILC. It is shown that the learning performance is significantly affected by the soft errors in a linear discrete time system.

3.1 Illustrative example of ILC

Figure 1 shows the basic structure of the ILC [1]. As shown in Figure 1 and equation (5), the input of the next iteration is determined from the output error and input of the current iteration. Since a large amount of memory should be used for the ILC, it can be easily affected by the soft errors in the memory.

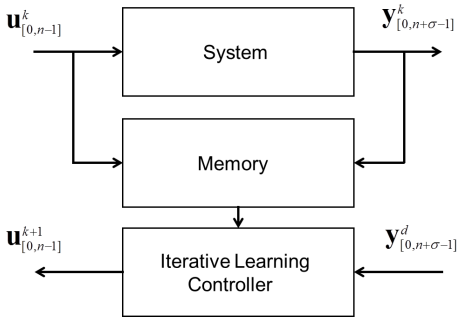


Figure 1: Basic ILC configuration for discrete time systems (modified from the figure in [1])

For the analysis of the effect of soft errors, let us consider an example of a linear discrete-time single-input and single-output(SISO) systems.

Let us consider the following linear discrete-time system:

$$\begin{aligned} \mathbf{x}(i+1) &= \begin{bmatrix} 0 & 0 & -0.1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \mathbf{x}(i) + \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} u(i) \\ y(i) &= \begin{bmatrix} 1 & 0.5 & 0.0625 \end{bmatrix} \mathbf{x}(i). \end{aligned} \tag{8}$$

This system has two minimum-phase zeros ($z = -0.25$) and the relative degree of the system is 1. The desired trajectory is given as follows:

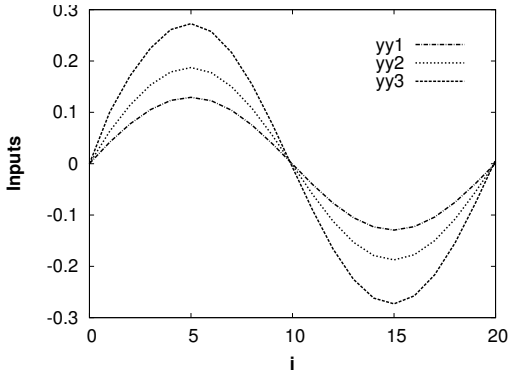
$$y^d(i) = \begin{cases} 0, & i = 0 \\ 0.2\sin(0.1\pi(i-1)), & 1 \le i \le 21. \end{cases} \tag{9}$$

Here, we set $N = 21$, $u(21) = 0$ and $\mathbf{S} = \alpha I_{21\times 21} = 0.1 I_{21\times 21}$. Since the system (8) and the learning gain α satisfy the conditions $||J_0|| > \sum_{i=1}^{N-1} ||J_i||$ with $1 > 0.7361$ and $|1 - \alpha J_0| < 1$ with $0.9 < 1$, which are sufficient conditions for the convergence in [16], the convergence condition $||I - 0.1\mathbf{J}|| < 1$ is satisfied.

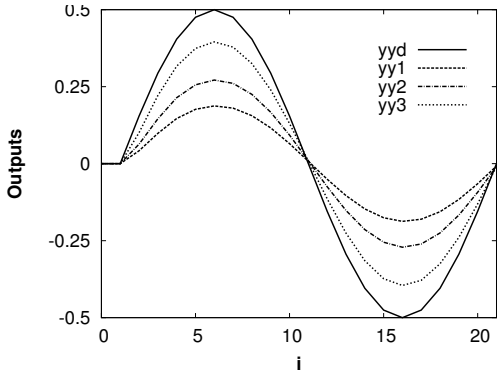
The input update law is expressed as

$$\mathbf{u}^{k+1}_{[0,20]} = \mathbf{u}^k_{[0,20]} + 0.1\mathbf{e}^k_{[1,21]}. \tag{10}$$

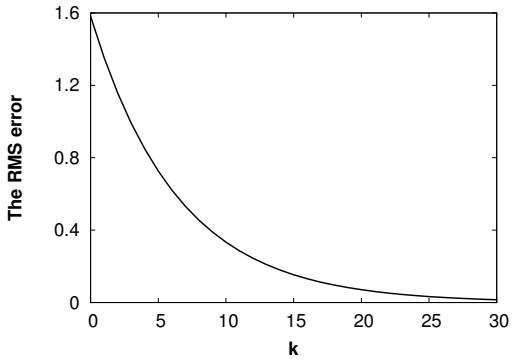
Figure 2 shows the outputs and inputs when there is no soft error. As shown in Figure 2(c), the root mean square(RMS) error for the output error approaches 0 as $k \to \infty$.



(a) Inputs for different values of k



(b) Desired output and the outputs for different values of k



(c) RMS error for different values of k

Figure 2: Outputs and inputs when there is no soft error

3.2 Effect of soft errors to ILC

Let us assume that a soft error occurs in the memory. If we assume that there is single bit error in the 4-byte floating point variable, the effect of the soft error is different, depending on the bit location among the 32 bits where a single soft error occurs. Table 1 summarized the structure of the 4-byte floating point variable [6].

Table 1: Binary format for floating point variable with single precision (32bits) [6]

Parts	Sign	Exponent	Fraction
Number of bits	1	8	23

Table 2 presents an example of the effect of the single soft error on the 4-byte floating point variable with a value of 0.2. As summarized in Tables 1 and 2, the soft errors in the sign bit, exponent part, and the upper part of the fraction lead to a large number of errors introduced and, eventually, can be critical to the control system. In real embedded systems, 1-bit error can be overcome by ECC(Error Correction Code). However when there are 3 or more bit errors, ECC cannot check those errors. Also the error values can be larger than that of 1-bit error. Assuming 1-bit error in MSB can show the effect of soft-errors in ILC. In order to show the effect of soft-errors, we assume just 1-bit error for the simulation without loss of generality.

By considering above mentioned aspects, let us discuss the effects of the soft error on the

Table 2: Effect of soft error on the 4-byte floating point variable with a value of 0.2

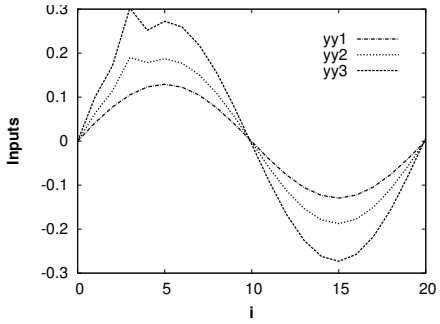
Part	Error bit	Approximate vaule (≈)
Sign	1(MSB)	-0.200000
Exponent	2	6.805647×10^{35}
	3	0.000000
	4	0.000000
	5	0.000003
	6	0.000781
	7	0.012500
	8	0.800000
	9	0.400000
Fraction	10	0.137500
	11	0.231250
	12	0.215625
	13	0.192188
	14	0.196094
	15	0.201953
	16	0.200977
	17	0.199512
	18	0.199776
	19	0.200122
	20	0.200061
	21	0.199969
	22	0.199985
	23	0.200008
	24	0.200004
	25	0.199998
	26	0.199999
	27-32	0.200000

ILC of the previous example for system (8). For the previous example, let us assume that there are soft errors in the memory. To analyze the effect of soft errors, we also assume the followings:

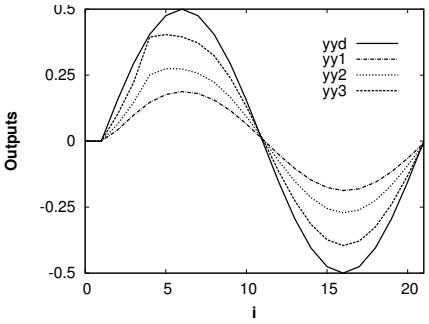
- 4-byte floating point variables are used for the storage and processing of all the data.
- 1-bit soft error occurs in $y^k(4)$ in the most significant bit(MSB) of the 4-byte memory.
- A soft error occurs at every 4-th iteration corresponding to the iteration number k.

With this setting as an example, since the MSB of the 4-byte floating number variable is the sign bit, the sign of $y^k(4)$ changes at every 4-th iteration. The learning is performed under these assumptions for the soft errors.

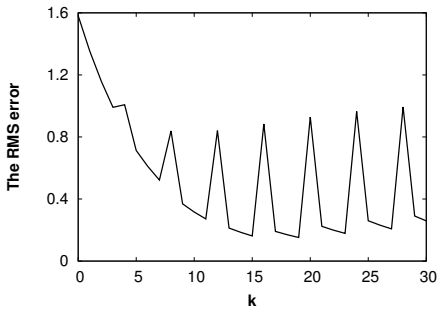
Figure 3 shows the effects of a soft error for the previous example. As shown in Figure 3, a soft error can significantly degrade the performance of the learning control. In other words, the soft errors reduce the learning speed and affect the convergence of the output error in the ILC.



(a) Inputs for different values of k



(b) Desired output and the outputs for different values of k



(c) RMS error for different values of k

Figure 3: Outputs and inputs when there are soft errors induced

4 Cross-layer approach to compensate for the soft error

In this section, we propose new detection and compensation methods for the soft errors in the ILC by obtaining the monotonic convergence with the cross-layer approach, as described in Figure 4. As summarized in Table 2, depending on the location of the soft error, its effect on the control system varies. The errors in the upper part of the 4-byte data format are critical to the control systems, while the other errors are negligible.

Similarly, with regard to the monotonic convergence of the output error, it is not required to detect all the soft errors. In this study, we detect the soft error that prohibits the monotonic convergence, and further, we compensate for the soft error with other relevant data available.

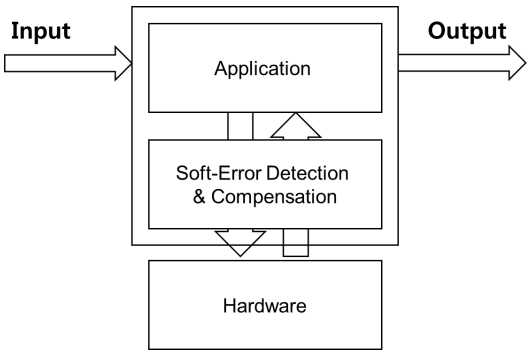


Figure 4: Compensation of soft error in a cross-layer manner

For further analysis, some assumptions are made by considering the soft error imposed on the system as follows:

- **(A1)** A soft error occurs in the system output, that is, y^k.
- **(A2)** At iteration k, only one soft error occurs. Hence soft-error occurs once at iteration k.
- **(A3)** There is no contiguous soft error corresponding to iteration k. In other words, when there is a soft error at iteration k, there is no soft error at $k+1$ and there may be a soft error at $k+2$.

With regard to (**A1**), soft errors can occur in the system, restored memories of the ILC, or the learning controller. Different from other variables, the system output y^k can have soft errors within all the components such as the system, restored memory, and learning controller. By considering these aspects, we make an assumption (**A1**).

Using (**A2**), we can detect the location of the soft error in the 4-byte floating point variable. Further, using (**A3**), we can derive the monotonic convergence with the output error of the previous iteration.

It should be noted that depending on the compensation method used, we can neglect or alleviate the related assumptions.

4.1 Detection of critical soft errors using cross-layer approach

If condition (6) is satisfied, $||\mathbf{e}^{k+1}_{[\sigma,N+\sigma-1]}|| \leq \rho||\mathbf{e}^{k}_{[\sigma,N+\sigma-1]}||$ should be guaranteed for some $\rho < 1$.

However, as shown in Figure 3, and Tables 1 and 2, depending on the location of the soft error, the monotonic convergence cannot be guaranteed.

Among all the soft errors, we only detect those errors affecting the monotonic convergence. Therefore, when $||\mathbf{e}^{k+1}_{[\sigma,N+\sigma-1]}|| \geq ||\mathbf{e}^{k}_{[\sigma,N+\sigma-1]}||$, we can conclude that a critical soft error has occurred. After detecting a soft error, we can find the location and compensate for the soft error.

Considering these facts, we propose a soft error detection algorithm to find the location of error j and iteration k for the ILC as follows :

Detection of critical soft errors in a cross-layered manner

- Step 1: For the learning curve, validate whether $||\mathbf{e}^k||_\infty \geq ||\mathbf{e}^{k-1}||_\infty$.
- Step 2: If the error $||\mathbf{e}^k||_\infty$ is larger than or equal to $||\mathbf{e}^{k-1}||_\infty$, it is determined that a critical soft error occurs at iteration k.
- Step 3: Determine j, which makes $||e^k(j)|| = ||\mathbf{e}^k||_\infty$, that is, the location of the largest value at iteration k.

The following lemma shows that we can detect the location of the soft error with the assumptions (**A1**)-(**A3**).

Lemma 2. *(Determining time j and iteration k, where soft error occurs.)*
*Let us assume that a soft error satisfying (**A1**)-(**A3**) occurs in the process of the ILC in the uncertain system (1). Using the above algorithm, the location of the soft error, that is, time j and iteration k can be determined.*

Proof: If there is no soft error, (7) holds for $\mathbf{e}^k$ and $\mathbf{e}^{k-1}$. However, if there is a critical soft error, we can obtain $||\mathbf{e}^k|| \geq ||\mathbf{e}^{k-1}||$. Thus, we can conclude that a critical soft error occurs at iteration k. □

When there is a critical soft error at iteration k, let us assume that the soft error occurs in j, that is, $e^k(j)$ contains the soft error.

Let us denote $\tilde{e}^k(j)$ as the real output error when there is no soft error, and $\tilde{\mathbf{e}}^k$ by replacing $e^k(j)$ with $\tilde{e}^k(j)$ in $\mathbf{e}^k$. From the definition of the ∞-norm, we can obtain the following equations:

$$\begin{aligned} &||\mathbf{e}^k||_\infty \\ &= \max\{||e^k(\sigma)||, \cdots, ||e^k(j)||, \cdots, ||e^k(N+\sigma-1)||\} \end{aligned} \tag{11}$$

$$\begin{aligned} &||\tilde{\mathbf{e}}^k||_\infty \\ &= \max\{||e^k(\sigma)||, \cdots, ||\tilde{e}^k(j)||, \cdots, ||e^k(N+\sigma-1)||\} \end{aligned} \tag{12}$$

Further, according to the relationship among $\mathbf{e}^k$, $\tilde{\mathbf{e}}^k$, and $\mathbf{e}^{k-1}$, the following equations hold:

$$||\mathbf{e}^k||_\infty \geq ||\mathbf{e}^{k-1}||_\infty, ||\tilde{\mathbf{e}}^k||_\infty < ||\mathbf{e}^{k-1}||_\infty. \tag{13}$$

Therefore, to satisfy (13), the following equation should be guaranteed by using (11) and (12).

$$||\mathbf{e}^k||_\infty = ||e^k(j)|| \tag{14}$$

Therefore, if we select the maximum value among the output errors at iteration k, we can determine j.

4.2 Compensation of soft errors

After detecting a soft error in $y^k(j)$, we can compensate for the soft error in several ways. For instance, as shown in Figure 5, we can use the stored data for $y^k(j)$, the value of the previous iteration $y^{k-1}(j)$, the adjacent value $y^k(j-1)$, or the desired output $y^d(j)$.

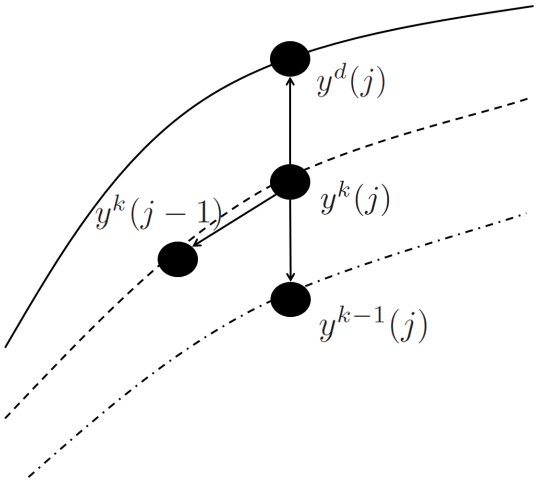


Figure 5: Recovery of the soft error induced $y^k(j)$ with related data

In general, for the resilient design of a soft error, the exact rollback method has been widely used. After periodically copying the value of $\mathbf{y}^k$ to the buffers, we can restore the value when an error is detected. Although we need additional buffer spaces for the rollback, it is a simple

and effective technique to reduce the memory errors. If advanced error-detection methods can be supported, it can be used without assumptions (**A2**) and (**A3**).

However, with regard to (**A1**), this method cannot be used when the soft error occurs within the system, that is, before copying the data to the buffer. For the soft errors in the memory or learning controller, the rollback method can be used irrespective of the variables.

The value of the previous iteration $y^{k-1}(j)$ can be used to compensate for $y^k(j)$. Since the learning is performed according to the iteration number k, it is easy to prove the monotonic convergence. In addition, it can be used when there is a soft error within the system or before copying the data to the buffer.

Further, we can use the adjacent value $y^k(j-1)$. For real-time control, this approach can be used for the compensation of soft errors. In this example, we do not need additional buffers. If more advanced error-detection methods can be supported, it can be used, relaxing assumptions (**A2**) and (**A3**).

Otherwise, the desired value $y^d(j)$ can be applied for the compensation. Since the error $e^k(j)$ becomes zero, monotonic convergence can be derived.

Based on these facts, we can summarize the compensation method as follows:

Compensation method 0 - exact rollback

- Step 1: Restore the measured output value $\mathbf{y}^k$ to an additional buffer at iteration k.
- Step 2: When a soft error is detected at iteration k and time j, recover it with $y^k(j)$ from the buffer.

Compensation method 1- using the value of the previous iteration $y^{k-1}(j)$

- Step 1: Restore the output value $\mathbf{y}^{k-1}$ to the memory at iteration $k-1$.
- Step 2: When a soft error is detected at iteration k and time j, replace $y^k(j)$ with $y^{k-1}(j)$.

Compensation method 2- using the adjacent value $y^k(j-1)$

- Step 1: When a soft error is detected at iteration k and time j, replace $y^k(j)$ with $y^k(j-1)$.

Compensation method 3- using the desired value $y^d(j)$

- Step 1: When a soft error is detected at iteration k and time j, replace $y^k(j)$ with $y^d(j)$.

With the presented compensation methods, monotonic convergence is proved in the theorems below. For compensation methods 1 and 3, convergence is given based on (**A1**)-(**A3**), whereas, for the compensation method 2, we need another condition for the magnitude of the compensated error.

Theorem 3. *(Convergence of the ILC when using compensation method 1)*
*Let us assume that a soft error satisfying (**A1**)-(**A3**) occurs at the process of the ILC in the uncertain system (1). Let us apply the input update law (5), the critical soft error detection method, and compensation method 1 to the uncertain system (1). If the condition (6) holds for all k, the output error $||e^k_{[\sigma,N+\sigma-1]}||_\infty$ converges to 0 as $k \to \infty$.*

Proof: Let us consider that a soft error occurs in $y^k(j)$. That is, $||\mathbf{e}^k||_\infty \geq ||\mathbf{e}^{k-1}||_\infty$ and $||\mathbf{e}^k||_\infty = ||e^k(j)||$.

From **Lemma 2**, we can detect j and k. Further, we compensate for the data using $y^{k-1}(j)$ instead of $y^k(j)$. Let us denote the compensating value $y^{k-1}(j)$ as $\hat{y}^k(j)$ and the compensated k-th iteration data as $\hat{\mathbf{e}}^k$, respectively. From these values, we obtain $\hat{\mathbf{e}}^k = [e^k(\sigma), \cdots, \hat{e}^k(j), \cdots, e^k(N+\sigma-1)]$, where $\hat{e}^k(j) = y^d(j) - \hat{y}^k(j) = y^d(j) - y^{k-1}(j) = e^{k-1}(j)$.

since $||\hat{e}^k(j)|| = ||e^{k-1}(j)||$, from (13) and (14), we can obtain

$$||\hat{\mathbf{e}}^k||_\infty \leq ||\mathbf{e}^{k-1}||_\infty. \tag{15}$$

Further, since there is no soft error at iteration $k+1$ from (**A3**), the following equation holds:

$$||\mathbf{e}^{k+1}||_\infty < \rho||\hat{\mathbf{e}}^k||_\infty, \rho < 1. \tag{16}$$

From (15) and (16), the output error approaches 0 as $k \to \infty$.

□

Theorem 4. *(Convergence of the ILC when using compensation method 2)*
*Let us assume that a soft error satisfying (**A1**)-(**A3**) occurs at the process of ILC in the uncertain system (1). Let us apply the input update law (5), the critical soft error detection method and compensation method 2 to the uncertain system (1). If the compensated error satisfies $||y^d(j) - y^k(j-1)|| \leq ||\mathbf{e}^{k-1}||_\infty$ and condition (6) holds for all k, the output error $||e^k_{[\sigma,N+\sigma-1]}||_\infty$ converges to 0 as $k \to \infty$.*

Proof: We can detect j and k with the detection method. Let us compensate for the data using $y^k(j-1)$ instead of $y^k(j)$. Let us denote the compensating value $y^k(j-1)$ as $\hat{y}^k(j)$ and the compensated k-th iteration error as $\hat{\mathbf{e}}^k$, respectively. From these values, we obtain $\hat{\mathbf{e}}^k = [e^k(\sigma), \cdots, \hat{e}^k(j), \cdots, e^k(N+\sigma-1)]$, where $\hat{e}^k(j) = y^d(j) - \hat{y}^k(j) = y^d(j) - y^k(j-1)$.

If $||\hat{e}^k(j)|| = ||y^d(j) - y^k(j-1)|| \leq ||\mathbf{e}^{k-1}||_\infty$, from (13) and (14), we can obtain

$$||\hat{\mathbf{e}}^k||_\infty \leq ||\mathbf{e}^{k-1}||_\infty. \tag{17}$$

Further, since there is no soft error at iteration $k+1$ from (**A3**), the following equation holds:

$$||\mathbf{e}^{k+1}||_\infty < \rho||\hat{\mathbf{e}}^k||_\infty, \rho < 1. \tag{18}$$

From (17) and (18), the output error approaches 0 as $k \to \infty$.

□

Theorem 5. *(Convergence of the ILC when using compensation method 3)*
*Let us assume that a soft error satisfying (**A1**)-(**A3**) occurs in the process of the ILC in the uncertain system (1). Let us apply the input update law (5), the critical soft error detection method and compensation method 3 to the uncertain system (1). If the condition (6) holds for all k, the output error $||e^k_{[\sigma,N+\sigma-1]}||_\infty$ converges to 0 as $k \to \infty$.*

Proof: We can detect j and k with the detection method. Let us compensate for the data using $y^d(j)$ instead of $y^k(j)$. Then, for the compensated data at iteration k, we obtain $\hat{\mathbf{e}}^k = [e^k(\sigma), \cdots, \hat{e}^k(j) = 0, \cdots, e^k(N+\sigma-1)]$.

Hence, from (13) and $\hat{e}^k(j) = 0$, we can obtain

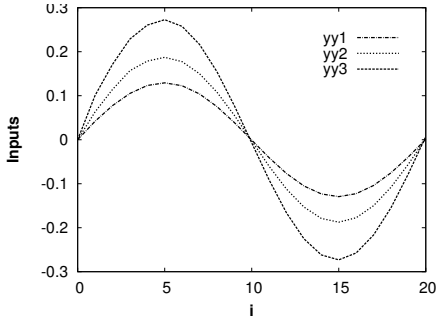
$$||\hat{\mathbf{e}}^k||_\infty \leq ||\tilde{\mathbf{e}}^k||_\infty < ||\mathbf{e}^{k-1}||_\infty. \tag{19}$$

Further, since there is no soft error at iteration $k+1$ from (**A3**), the following equation holds:

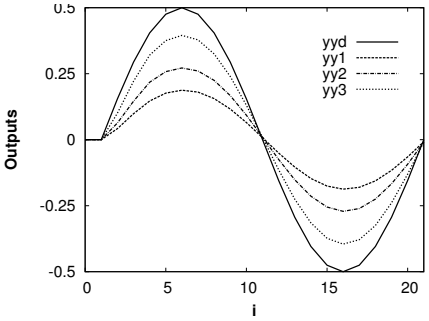
$$||\mathbf{e}^{k+1}||_\infty < \rho||\hat{\mathbf{e}}^k||_\infty, \rho < 1. \tag{20}$$

From (19) and (20), the output error approaches 0 as $k \to \infty$.

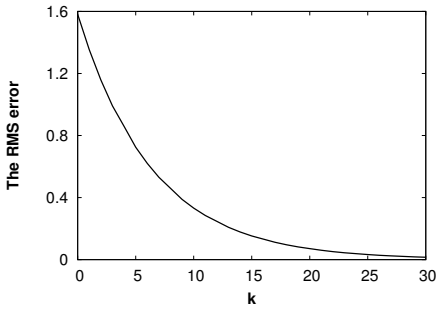
□



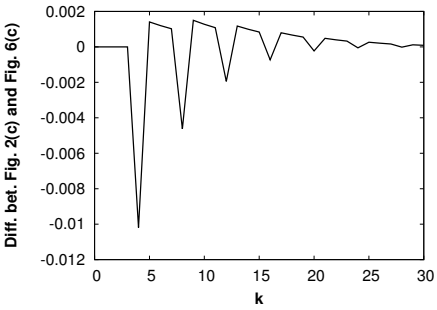
(a) Inputs for different values of k



(b) Desired output and the outputs for different values of k



(c) RMS error for different values of k



(d) Differences between the output errors when there is no soft error and when it is compensated using method 1

Figure 6: Outputs and inputs when soft error compensation method 1 is used

5 Simulation results

Let us apply the proposed methods to the case in Section 3.2. The soft error condition in the previous example satisfies assumptions (**A1**)-(**A3**) and we can apply the detection and compensation methods to this example.

Figure 6 shows the results of compensation method 1. As shown in Figure 6, we can observe that the output error converges to 0 as $k \to \infty$ and the compensation method works well although there is a slight performance degradation as compared with the case without soft errors. Figure 6(d) shows the differences between the output errors when there is no soft error and when the soft error is compensated using compensation method 1.

Figure 7 shows $||\mathbf{e}^k||_\infty$ trajectory when there is no soft error and when the induced soft error is compensated using the compensation methods. For compensation methods 1 and 3, we can observe that the error approaches 0 as $k \to \infty$, whereas for compensation methods 2, $||y^d(j) - y^k(j-1)|| \leq ||\mathbf{e}^{k-1}||_\infty$ cannot always be satisfied, and the monotonic convergence cannot be guaranteed.

In this example, when using compensation method 3, as compared to compensation method 1, $\hat{e}^k(j)$ becomes 0, and it reduces the speed of convergence. As shown in Figure 7, compensation method 1 is the best method in this example, among compensation methods 1, 2, and 3, considering the monotonic convergence in the ILC. Further, depending on the location and characteristics of the soft errors, we can combine the presented four compensation methods for achieving a better performance.

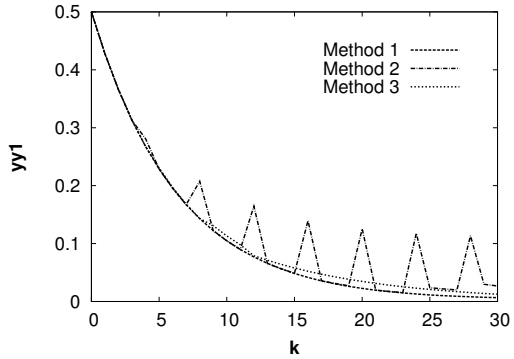


Figure 7: $||\mathbf{e}^k||_\infty$ according to k for the different compensation methods

6 Conclusion

In this paper, we have studied the effects of soft errors in the ILC and proposed novel compensation methods for the soft errors in the ILC. First, we have analyzed the effects of soft errors in the ILC. It has been shown that a soft error can affect the speed of the learning curve and make the controller lose the learning history. Second, based on the detection algorithm for the critical soft error that affects the monotonic convergence, we have proposed compensation methods for the ILC. The convergence of the proposed method has been proved under the assumptions for soft errors. From the experimental results, we have demonstrated the effectiveness of our proposed methods.

Furthermore, it is expected that several directions of researches could be developed. We could make more complicated assumptions such as multiple soft errors and extend the results to the various control applications. As an example in the ILC, depending on the location and characteristics of the soft errors, we could combine the four presented compensation methods for achieving better performance. These remain as future work.

Funding

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2018R1D1A1A09083894), and also supported by the National Research Foundation of Korea (NRF) funded by the Korean Government (MSIP) (NRF-2016R1A5A1012966).

Author contributions

The authors contributed equally to this work.

Conflict of interest

The authors declare no conflict of interest.

Bibliography

[1] Ahn, H.-S.; Chen, Y.; Moore, K. L. (2007). Iterative Learning Control: Brief Survey and Categorization, *IEEE Trans. Syst., Man and Cybern., Part C,* 37(6), 1099-1121, 2007.

[2] Arimoto, S.; Kawamura, S.; Miyazaki, F. (1984). Bettering operation of robots by learning, *J. Robotic Systems*, 1(2), 123–140, 1984.

[3] Azizimazreah, A., et al. (2018). Tolerating soft errors in deep learning accelerators with reliable on-chip memory designs, *2018 IEEE International Conference on Networking, Architecture and Storage (NAS)*, 1-10, 2018.

[4] Baumann, R. (2005). Soft errors in advanced computer systems, *Design & Test of Computers*, 22(3), 258-266, 2005.

[5] Bien, Z.; Xu, J.-X. (1998). *Iterative Learning Control Analysis, Design, Integration and Applications*, Kluwer Academic Publishers, 1998.

[6] IEEE Computer Society, *IEEE Standard for Floating-Point Arithmetic(IEEE Std 754)*, IEEE, 2008.

[7] Jeong, G.-M.; Choi, C.-H. (2002). Iterative learning control for linear discrete time nonminimum phase systems, *Automatica*, 38(2), 287–291, 2002.

[8] Jeong, G.-M., et al. (2016). Robust face recognition against soft-errors using a cross-layer approach, *International Journal of Computers Communications & Control*, 11(5), 657-665, 2016.

[9] Jeong, G.-M.; Ji, S.-H. (2017). Learning speed enhancement of iterative learning control with advanced output data based on parameter estimation, *International Journal of Computers Communications & Control*, 12(3), 323-329, 2017.

[10] Khudia,D. S.; Wright, G.; Mahlke, S. (2012). Efficient soft error protection for commodity embedded microprocessors using profile information, In *the 13th ACM SIGPLAN/SIGBED International Conference on Languages, Compilers, Tools and Theory for Embedded Systems (LCTES '12)*, 99–108, 2012.

[11] Lee, K., et al. (2006). Mitigating soft error failures for multimedia applications by selective data protection, *Proceedings of the 2006 international conference on Compilers, architecture and synthesis for embedded systems*, 411-420, 2006.

[12] Lee, K., et al. (2008). Mitigating the impact of hardware defects on multimedia applications: a cross-layer approach, *In: Proceedings of the 16th ACM international conference on Multimedia*, 319-328, 2008.

[13] Lee, K., et al. (2007). Partially protected caches to reduce failures due to soft errors in multimedia applications, *IEEE Transactions on Very Large Scale Integration (VLSI) Systems*, 17(9), 1343-1347, 2007.

[14] Lyons, D. (2000). *Sun screen: Soft error issue in sun enterprise servers*, Forbes, 2000.

[15] Michalak, S. E., et al.(2005). Predicting the number of fatal soft errors in los alamos national laboratory's ASC Q supercomputer, *IEEE Transactions on Device and Materials Reliability*, 5(3), 329-335, 2005.

[16] Moore, K. L.; Chen, Y.; Bahl, V. (2005). Monotonocally convergent iterative learning control for linear discret time systems, *Automatica*, 41(1), 1529–1537, 2005.

[17] Mukherjee, S. (2008). *Architecture Design for Soft Errors*, Morgan Kaufmann Publishers Inc., 2008.

[18] Nakka, N.; Pattabiraman, K.; Iyer, R. (2007). Processor-Level Selective Replication, *In: Proceedings of the 37th Annual IEEE/IFIP International Conference on Dependable Systems and Networks*, 544-553, 2007.

[19] Tremblay, M.; Tami, Y. (1989). Support for fault tolerance in VLSI processors, *IEEE International Symposium on Circuits and Systems*, 388-392, 1989.

[20] Uchiyama, M. (1978). Formulation of high-speed motion pattern of mechanical arm by trial, *Trans. Soc. Instrum. and Contr. Eng.* (in Japanese), 14(6), 706–712, 1978.

[21] Vera, X., et al. (2009). Selective replication: A lightweight technique for soft errors, *ACM Transactions on Computer Systems*, 27(4), 1-30, 2009.

[22] Wang, N. J.; Quek, J.; Rafacz, T.; Patel, S. (2004). Characterizing the effects of transient faults on a high-performance processor pipeline, *2004 Intl. Conf. on Dependable Systems and Networks*, 61-70, 2004.

INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL
ISSN 1841-9836, e-ISSN 1841-9844, 14(3), 375-387, June 2019.

Wavelet Design for Automatic Real-Time Eye Blink Detection and Recognition in EEG Signals

M. Miranda, R. Salinas, U. Raff, O. Magna

Michael Miranda*
Department of Informatic Engineering
Metropolitan University of Technology, Chile
Jose Pedro Alessandri 1242, Nunoa, Santiago, Chile
*Corresponding author: michael.miranda@utem.cl

Renato Salinas
Department of Mechanical Engineering
University of Santiago, Chile
Av. Libertador Bernardo O'Higgins 3363, Santiago, Chile.
renato.salinas@usach.cl

Ulrich Raff
Department of Physics
University of Santiago, Chile
Av. Libertador Bernardo O'Higgins 3363, Santiago, Chile.
ulrich.raff@usach.cl

Oscar Magna
Department of Informatic Engineering
Metropolitan University of Technology, Chile
Jose Pedro Alessandri 1242, Nunoa, Santiago, Chile
omagna@utem.cl

Abstract: The blinking of an eye can be detected in electroencephalographic (EEG) recordings and can be understood as a useful control signal in some information processing tasks. The detection of a specific pattern associated with the blinking of an eye in real time using EEG signals of a single channel has been analyzed. This study considers both theoretical and practical principles enabling the design and implementation of a system capable of precise real-time detection of eye blinks within the EEG signal. This signal or pattern is subject to considerable scale changes and multiple incidences. In our proposed approach, a new wavelet was designed to improve the detection and localization of the eye blinking signal. The detection of multiple occurrences of the blinking perturbation in the recordings performed in real-time operation is achieved with a window giving a time-limited projection of an ongoing analysis of the sampled EEG signal.
Keywords: Biological signals, electroencephalogram, brain computer interface, eye blink detection, pattern recognition, wavelet design.

1 Introduction

The electroencephalogram was designed to record the brain activity of a living being, specifically by sensors, with which electrical signals from the brain are captured. Any activity that does not come from the brain, that is, noise, is called an artifact. Depending on the origin, they can be divided into physiological and extra-physiological artifacts. The first ones are generated by the patient, on the other hand, extra-physiological artifacts arise from sources external to the individual, that is to say from the environment. Blinking of eyes is considered an ocular artifact, which is captured in the frontal zones Fp1-Fp2 of the international system 10-20 [10].

The eyeball acts as a dipole, with a positive pole oriented anteriorly (cornea) and a negative pole oriented posteriorly (retina). A blinking of eyes causes the positive pole to approach the fronto-polar electrodes Fp1-Fp2, producing symmetrical descending deflections [15]. The EEG records are manifested with periodic and unpredictable oscillations, having a greater spectral amplitude in certain frequency bands, which are divided into: $0.5 - 4$ $Hertz$ (delta band), $4 - 8$ $Hertz$ (theta band), $8 - 12$ $Hertz$ (alpha band), $12 - 30$ $Hertz$ (beta band) and greater than 30 $Hertz$ (gamma band) [4].

The extraction of characteristics is an important process in the classification of EEG signals, and additionally, it is also possible to relate patterns with a defined intentionality on the part of the individual. However, the relationship between the electrical signal and the intention is diffuse, besides it is necessary to filter the artifacts, to obtain a signal without noise, so as not to disturb the results. One of the artifacts that must be filtered is the eye blinking, but before filtering it must be detected. The blinking of eyes could be considered, in itself, a control signal, since in general an individual can generate voluntary blinks, being able to increase the frequency and / or amplitude, which in turn can be associated with machine instructions for a computer. From the previous conceptualization, it is possible to devise a brain-machine interface or BCI, which uses eye blinking as support data. A brain-machine interface or BCI, is defined as the technology that allows capturing brain waves to be processed by a computer with the intention of obtaining information about a state or cognitive process of the person. For example, in Dzitac et al. [7] changes in amplitude are studied in the different frequency bands, specifically the events of desynchronization (or event-related ERD) and their application to BCI.

The work that will be presented in this publication, shows the research results related to the automatic detection of eye blinking by designing an ad-hoc mother wavelet, which allows the analysis in the time domain.

2 Materials and methods

2.1 Creation of an EEG signal database

The design of a mother wavelet requires experimental data and validation, from which it is possible to extract the representative pattern present in the study signal. For this reason, a free access database of EEG signals was consulted to obtain experimental data. Subsequently, a software based on Matlab and Java was developed for obtaining in real time experimental validation samples, recording the EEG signal and the video of the face of the individual with the same time-base, this way to correlate the video with the EEG signal and clearly define the occurrence of the blinks.

2.2 Obtaining experimental data

To verify the main hypotheses of this work, the available data is located in www.physionet.org, specifically the PhysioBank database. This is a database of high-growth physiological digital signals, with good foundation in data related to the biomedical research community. Polysomnographic record signals, of multiple parameters, are included in this database including cardiopulmonary, neural and other fields of biomedicine. The records correspond to both healthy individuals and patients with a wide variety of conditions that concern different pathological implications, such as sudden cardiac death, failure congestive heart, epilepsy, motor disorders, sleep apnea and senility. This database is freely accessible via web and tends to cooperative activities, that is providing data for research and requesting the submission of results for its feedback. The PhysioBank collections are organized in more than 50 databases, each with a number of records, and

each record contains information collected from a single subject.

2.3 Obtaining validation data

Validation data were obtained from a continuous EEG measurement with the biosensor consisting of a band designed by Neurosky MindWave TM, being one of the first to enter the market of EEG amplifiers for use in non-invasive BCI. Mindwave is very economical due to the simplicity that characterizes it, having one electrode in position FP1 and has the option of being compatible with both iOs and Android [13].

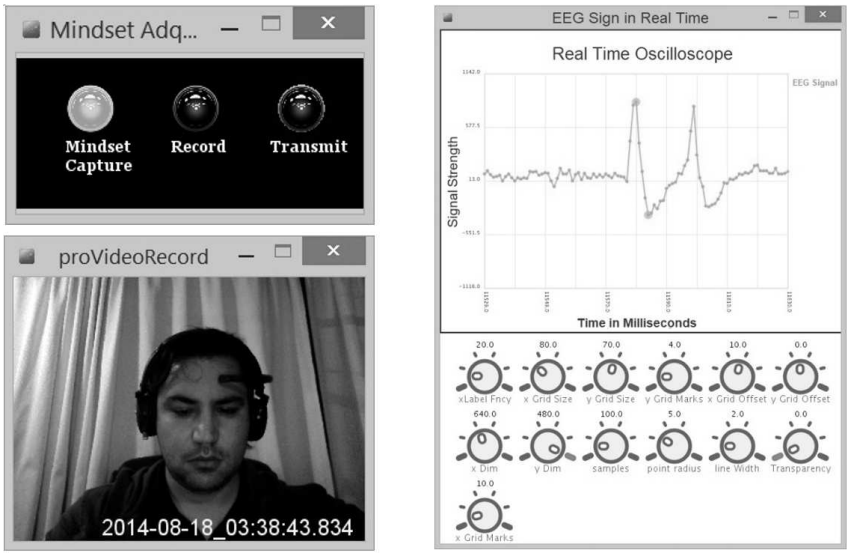


Figure 1: Video recording module, status indicator module and real-time software-oscilloscope module

The state of the eye blink was detected through a video camera during the EEG measurement, subsequently and then manually added to the file after analyzing the video frames. Where "1" indicates the closed eye state and "0" the state of the eye open. All the values are in chronological order with the first value measured in the top of the data. The platform developed is called "EEG Studio" and is composed of three modules:

1. Oscilloscope that works in real time.
2. Video recording: records the face of the individual in real time.
3. Status Indicator: indicates the correct functioning of the different modules.

3 Wavelet transform

Wavelets are functions that satisfy certain mathematical requirements and are used for the representation of data or other functions. Wavelets are very suitable for data approximation of signals with abrupt discontinuities. The fundamental idea behind wavelets is to analyze functions according to scales. In wavelet analysis, the scale used to analyze the data plays a special role. Algorithms that use wavelets process data at different scales or resolutions. If a signal or function is observed using a wide "window", the small details are not observed; On the other hand, if the "window" used is narrow, then they can be observed. In wavelet analysis,

these windows are automatically adjusted when changing resolution, usually referred to as multi-resolution analysis. This makes wavelets a useful and interesting tool. The general procedure of the analysis using wavelets is to adopt a "prototype" function, generally called mother wavelet. The temporal analysis is then carried out using dilatations and translations of said function. The original signal can then be represented as a linear combination of the original function and its translated and dilated ones. This procedure is called a wavelet expansion. The choice of the mother wavelet [13] (and thus the base or the wavelet frame) is not unique and depends on the type of functions or data to analyze. The multi-resolution analysis of the wavelets, makes it a very powerful tool for the study of EEG signals. An example of the EEG artifact and pattern of blinking of eyes is seen in Fig. 2.

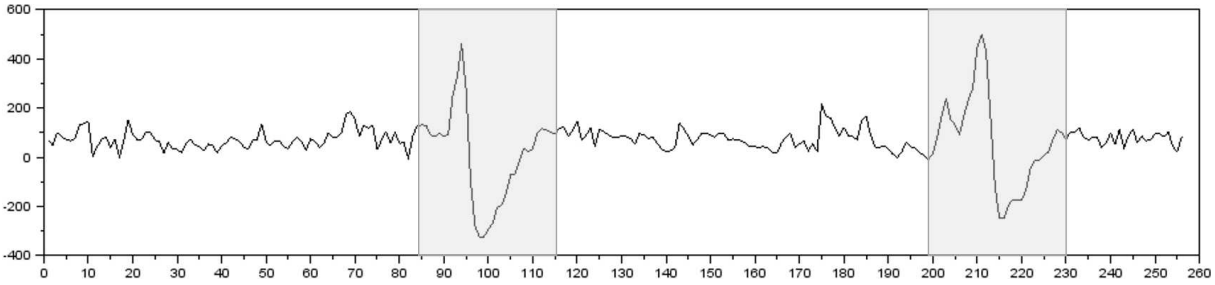


Figure 2: EEG signal with two artifacts produced by blinking eyes. The artifacts are indicated with two light gray time intervals

3.1 Continuous wavelet transform

The continuous Wavelet transform allows the analysis of a signal in a segment located in it and consists in expressing a continuous signal as an expansion of terms or coefficients of the internal product between the signal and a Mother Wavelet function. A Mother Wavelet is a localized function, belonging to space $L^2(R)$, that contains all functions with finite energy and square integrable functions defined:

$$f \in L^2 \Rightarrow \int |f(t)|^2 \, dt = E < \infty \tag{1}$$

In this way we have a single modulated window and from this a complete family of elementary functions is generated by dilatations or contractions and translations in time $\Psi_{a,b}(t)$, called wavelet daughters that meet all the conditions of the form:

$$\Psi_{a,b}(t) = \frac{1}{\sqrt{a}} \Psi \left(\frac{t-b}{a} \right) \qquad (a,b) \in R, \ a \neq 0. \tag{2}$$

The Mother Wavelet must comply with the eligibility condition:

$$C_\Psi(t) = \int_0^\infty \frac{|\Psi(\omega)|^2}{\omega} d\omega < \infty \tag{3}$$

Which means that the function is well localized in time, that is, the function oscillates around an axis and its average is zero and that the Fourier transform is a continuous band-pass filter, with rapid decrease towards infinity and toward $\omega = 0$. The Wavelet transform of a function at a given scale and position, is calculated by the correlation of the form:

$$CWT\,(a,b) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(t) \Psi \left(\frac{t-b}{a} \right) dt \tag{4}$$

3.2 Wavelet pattern design applied in CWT

When searching a defined pattern it is possible to apply two different techniques to perform multiresolution analysis. The first technique includes the use of an existent wavelet family trying to find whatever fits best the detection of the pattern, i.e. trial and error technique [16]. In reference [16] the biorthogonal wavelet was used. The other alternative consists in designing or building a wavelet starting with the pattern that is asked to be detected. This process can obviously decrease the time to approach the desired goal while improving the detection of the blinking artifact. This approach will work as long as the characteristics of the designed wavelet are close enough to the characteristics of the recorded pattern. Based on a given function f with compact support and finite energy, we may consider the construction of a wavelet usable with CWT while approximating this function in the least square sense. Various construction methods are possible, and since we have numerical sampling of our pattern over a given interval $[a, b]$, we used the method described by Misiti [14]. Let us consider a finite set of values:

$$(t_k, y_k)_{k=1,\,\ldots\,,\,K} \quad , \quad such\ that: \quad a \leq\ t_k \leq\ b\ and\ y_k \approx\ f(t_k) \tag{5}$$

Consider a family $F = \rho_i{}_{i=1}^{N}$ of linearly independent functions in $L^2(a,b)$, where L^2 is the space of square integrable functions over $\mathbb{R}$, and denote by V the vector space spanning F. Formulated for this finite set of pairs, the problem consists of seeking coefficients $\alpha = \alpha_i{}_{i=1}^{N}\ in\ \ \mathbb{R}^N$, where:

$$\psi = \sum_{i=1}^{N} \alpha_i \rho_i \tag{6}$$

such that:

$$\sum_{k=1}^{K} \left[\psi\,(t_k) - y_k \right]^2 = \underset{\beta \epsilon \mathbb{R}^N}{Min} \left\{ \sum_{k=1}^{K} \left[v_\beta\,(t_k) - y_k \right]^2 \right\} \tag{7}$$

such that:

$$\int_a^b v_\beta\,(t)\,dt = 0 \tag{8}$$

where for β in $\mathbb{R}^N$:

$$v_\beta = \sum_{i=1}^{N} \beta_i \rho_i \tag{9}$$

It is thus a problem of least squares sense minimization with a constraint. The vector α and the Lagrange multiplier λ associated with the constraint are obtained by solving the linear system:

$$\begin{bmatrix} G & M^t \\ M & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ \lambda \end{bmatrix} = \begin{bmatrix} B \\ 0 \end{bmatrix} \tag{10}$$

with G, M and B defined by:

$$G_{i,j} = \sum_{k=1}^{K} \rho_i\,(t_k) \cdot \rho_j\,(t_k) \tag{11}$$

$$M_i = \int_a^b \rho_i\,(t)\,dt \tag{12}$$

$$B_i = \sum_{k=1}^{K} y_k \cdot \rho_j\,(t_k) \tag{13}$$

The wavelet $\psi(t)$, must satisfy conditions of admissibility and regularity. Regularity is defined as the capability of a given wavelet to reconstruct a signal from the coefficients computed during the transformation process [11]. In other words, a function is regular if it can be locally approximated by a polynomial. According to Lipschitz's definition of regularity [12], a function f is pointwise Lipschitz $\alpha \geq 0$ at υ, if there exist $C > 0$, and a polynomial ρ_υ of degree $m = [\alpha]$, such that:

$$\forall t\ \epsilon\ \mathbb{R}\ \ ,\ \ |f\,(t) - \rho_\upsilon(t)| \leq C \cdot |t - \upsilon|^\alpha \tag{14}$$

A function f is uniformely Lipschitz α over $[a, b]$ if it satisfies for all $\upsilon\ \epsilon\ [a, b]$, with a constant C that is independent of υ. The Lipschitz regularity of f at υ or over $[a, b]$ is the supremum of the α such that f is Lipschitz α. On the other hand, admissibility [11] is defined by the following conditions:

$$\int_{-\infty}^{\infty} \psi\,(t)\,dt = 0 \tag{15}$$

$$\int_{-\infty}^{\infty} |\psi\,(t)\,|^2 dt < \infty \tag{16}$$

That is, function $\psi(t)$ must be localized in a bounded time interval, having oscillations around time axis, so its average be zero.

3.3 Characteristic pattern detection

Given the waveform pattern associated with eye blinking shown in Figure 2, it is necessary to design a mother wavelet as close as possible to the recorded pattern. This new wavelet will be labeled "blinkwave". This designed wavelet can then be used in the CWT analysis because it fulfills the required wavelet properties using the CWT technique [6].

4 Design and use of the mother wavelet

To design the new wavelet ("blinkwave") it becomes necessary to isolate the EEG Eye Blinking Pattern. Using this procedure, we obtain a vector or a set of finite number of values in equation (6). This vector is utilized in the process described by Misiti [14]. Designing the wavelet process with polynomials of variable grades, we obtain a mother wavelet with best approximation to the recorded pattern with a polynomial of grade 6, which can be observed in Figure 3.

In this case, the family $F = \rho_i{}_{i=1}^{N}$ used in equations (5) to (13), is the polynomial family of grade $N \geq 6$.

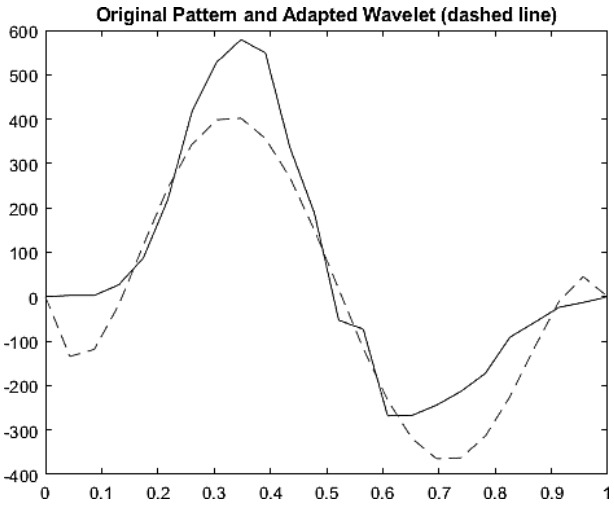


Figure 3: Overlay of the waveform produced by eye blinking and the wavelet designed, both normalized in the range $[-1, 1]$

4.1 Design of the mother wavelet

According to the method hinted in [18], it is necessary to construct matrices G, M and B, to fulfill equation (10), and to resolve the following linear system:

$$\begin{pmatrix} G & M^t \\ M & 0 \end{pmatrix} \begin{pmatrix} \alpha \\ \lambda \end{pmatrix} = \begin{pmatrix} B \\ 0 \end{pmatrix} \qquad \Rightarrow \qquad A \cdot U = B \tag{17}$$

I.e., solve U as a function of A:

$$U = A^{-1} \cdot B \tag{18}$$

We take as base the pattern ("blinkwave"), which is defined as an n-component vector:

$$\{y_i\}_{i=1}^{m} = y_1, \;\; y_2, \;\; \ldots, \;\; y_m \tag{19}$$

The base pattern is obtained digitizing the continuous signal with a constant sampling frequency f_s which defines the time interval Δt between samples. The matrix or file vector M, is built using the formula in equation (12), considering the admissibility restrictions given by equations (15) and (16), resulting in the following series:

$$M_{i=[1,m]} = \int_a^b \rho_i(t)\, dt = \left\{ \frac{1}{i}(b^i - a^i) \right\}_{i=[1,m]} =$$

$$= \left[(b-a), \;\; \left(\frac{b^2}{2} - \frac{a^2}{2} \right), \;\; \ldots \;\; , \left(\frac{b^{N+1}}{N+1} - \frac{a^{N+1}}{N+1} \right) \right] \tag{20}$$

If one chooses initiating the analysis at the origin, $a = 0$, then all computations become simpler, then, recomputing M we obtain:

$$M = \left[b, \;\; \left(\frac{b^2}{2} \right), \;\; \ldots \;\; , \left(\frac{b^{N+1}}{N+1} \right) \right]_{1 \times (N+1)} \tag{21}$$

Given M, we must apply the regularity condition indicated in equation (14), to get a new matrix M, as follows:

$$M = \begin{bmatrix} b, & \left(\frac{b^2}{2}\right), & \dots & , \left(\frac{b^{N+1}}{N+1}\right) \\ 1 & 0 & \dots & 0 \\ 1 & b & \dots & b^N \end{bmatrix}_{3\times(N+1)} \tag{22}$$

Afterwards, the matrix G is built using equation (11):

$$G = \begin{bmatrix} b & \left(\frac{b^2}{2}\right) & \cdots & \left(\frac{b^{N+1}}{N+1}\right) \\ \left(\frac{b^2}{2}\right) & \left(\frac{b^3}{3}\right) & \ddots & \left(\frac{b^{N+2}}{N+2}\right) \\ \vdots & \vdots & \cdots & \vdots \\ \left(\frac{b^{N+1}}{N+1}\right) & \left(\frac{b^{N+2}}{N+2}\right) & \cdots & \left(\frac{b^{2N+1}}{N+2}\right) \end{bmatrix}_{(N+1)\times(N+1)} \tag{23}$$

Prior to solving the linear system for U, we must generate the column vector B, as described in equation (13):

$$B = \left(\frac{1}{2}\right) \begin{bmatrix} \sum_{i=1}^{m} (x_{i+1} - x_i) \left(x_{i+1} \cdot y_{i+1} + x_i \cdot y_i\right) \\ \sum_{i=1}^{m} (x_{i+1} - x_i)({x_{i+1}}^2 \cdot y_{i+1} + {x_i}^2 \cdot y_i) \\ \vdots \\ \sum_{i=1}^{m} (x_{i+1} - x_i) \left({x_{i+1}}^N \cdot y_{i+1} + {x_i}^N \cdot y_i\right) \end{bmatrix}_{(N+1)\times 1} \tag{24}$$

Once matrices G and M are generated, it is possible assembling matrix A and proceed solving equation (17):

$$U = A^{-1} \cdot B \qquad \Rightarrow \qquad \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \lambda_1 \end{bmatrix} = \begin{bmatrix} G & M^t \\ M & 0 \end{bmatrix}^{.1} \cdot [B] \tag{25}$$

Mother wavelet $\psi(t)$, is generated using formula in equation (6) replacing the values obtained for α_i

$$\psi = \sum_{i=1}^{N} \alpha_i \cdot \rho_i \qquad \Rightarrow \qquad \psi(t) = \sum_{i=1}^{N+1} \alpha_i \cdot {x_i}^{N+1-i} \tag{26}$$

Thus, for the numerical pattern shown in Figure 3 (solid line), represented by vector:

$$y = [0.0, 1.7, 2.0, 1.6, 0.8, 0.0, -0.6, -0.9, -1.1, -1.0, -0.9, -0.7, -0.5, -0.4, -0.3, 0.0] \tag{27}$$

We obtain the polynomial mother wavelet $\psi(t)$ as follow:

$$\psi(t) = \alpha_1 \cdot x^n + \alpha_2 \cdot x^{n-1} + \ldots + \alpha_n \cdot x^n + \alpha_{n+1} \tag{28}$$

I.e., the first n numerical coefficients α_i are employed, obtained from solving equation (25) and discarding the Lagrange multipliers λ_i. These coefficients for the mother wavelet $\psi(t)$, for the pattern given in equation (27), are shown in equation (29) as follows:

$$p = [380, \quad -1.2 \cdot 10^3, \quad 1.3 \cdot 10^3, \quad -640,110, \quad 1.2, \quad -2.5 \cdot 10^{-14}] \tag{29}$$

Finally, the polynomial equation for the new mother wavelet $\psi(t)$, is shown graphically in Figure 3 (dashed line) and it is represented algebraically by equation (30), as follows:

$$\psi(t) = 380 \cdot x^6 - 1200 \cdot x^5 + 1300 \cdot x^4 - 640 \cdot x^3 + 110 \cdot x^2 + 1.2 \cdot x - 02.5 \cdot 10^{-14} \tag{30}$$

4.2 Use of the mother wavelet

The new wavelet-blinkwave is used to obtain the coefficients, which will subsequently be processed to obtain the location of each artifact. Using the signal shown in Figure 3 or the "blinkwave" mother wavelet for analysis and applying the CWT, we obtain a coefficient matrix $C(a,b)$ using equation (31), where:

$$C(a,b) \;=\; CWT(\text{"}inputsignal\text{"},\text{"}blinkwave\text{"}) \tag{31}$$

$C(a,b)$ is the coefficients matrix and CWT is the continuous wavelet transform, applied to input signal. The next step includes the detection and duration of the eye blinking from $C(a,b)$. To improve the analysis, a threshold must be applied to the matrix of wavelet coefficients, as defined in equation (32).

$$T(a,b) = \begin{cases} 0 & if \quad C(a,b) < Threshold_{fixed} \\ C(a,b) & , \qquad otherwise \end{cases} \tag{32}$$

The dimensions of the matrix depend on the number of samples of the analyzed signal and the level or number of scales in the decomposition performed by the wavelet transform.

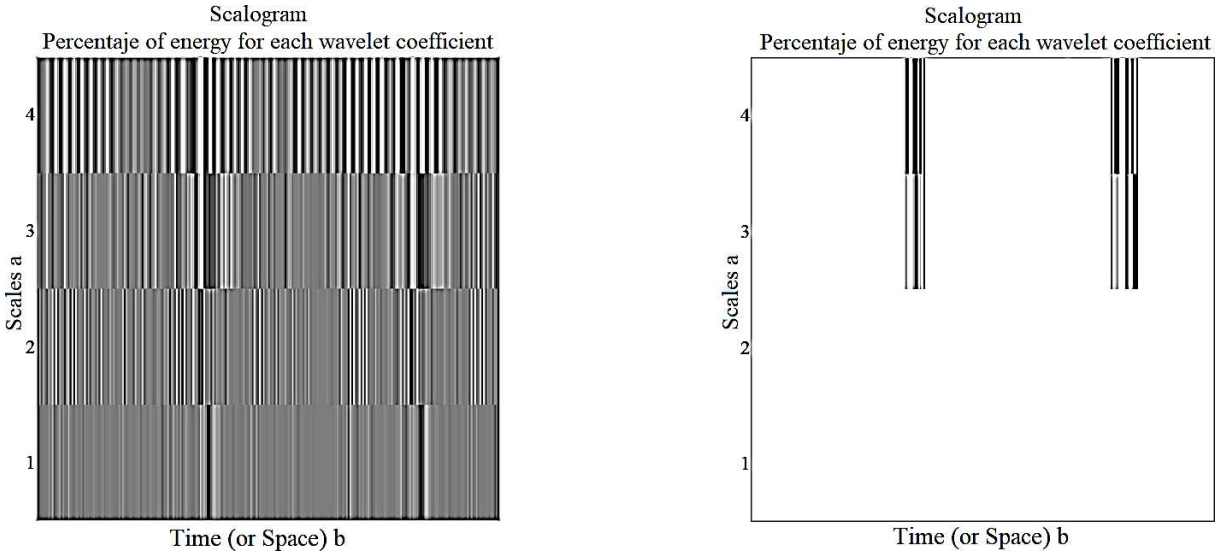


Figure 4: Graphical representation or scalogram to the left image and the thresholded wavelet coefficient matrix $T(a,b)$ to the right

As can be observed in Figure 4 there is a defined relation between the coefficients and the searched pattern. In our case, the graphical representations displayed in Figure 4 are used to support the detection of two blinking artifacts in the EEG recording of Figure 2. A threshold has been applied to the matrix of wavelet coefficients, as defined in equation (14). With this numerical relation, it is possible to determine the temporal location and duration of the pattern in the signal analysis process. In the case studied in this work, the number of samples b is

256 (analysis window size) and 4 levels of decomposition, which involves obtaining a coefficient matrix $[C]$ of $4x256$ elements. To improve the location of the blinking patterns, the columns of the coefficient matrix $T(a,b)$ were added as shown in equation (33).

$$V(1,2,\ldots,\ n)\ = \sum_{a} T\,(a,b)\,, \quad b = 1,2,\ldots,n \tag{33}$$

The vector V obtained gives us the actual location in time of the artifact produced by the eye blinking. To achieve a tradeoff between data smoothing and real time analysis, a moving average of a small number of samples with a range of $10[ms]$ was applied. Applying the system shown in Figure 1 to the recorded signal displayed in Figure 2 allows overlying the dotted signal shown in Figure 5, which detects the occurrence of eye blinking and their duration.

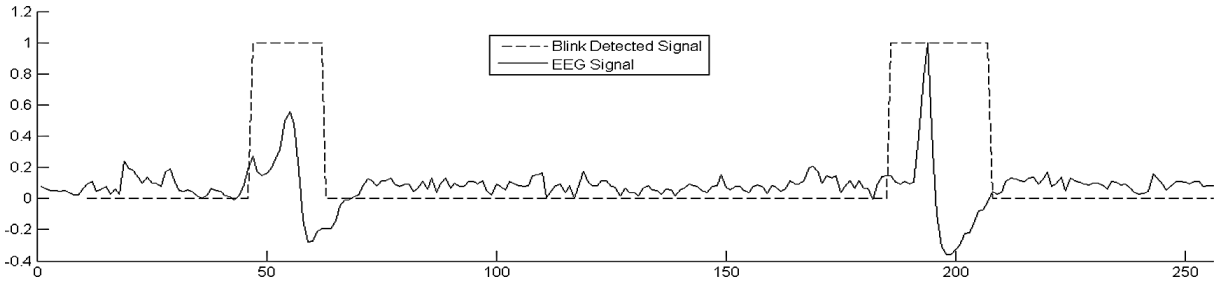


Figure 5: Overlay of original EEG signal with the dotted signal highlighting the detected artifacts and their duration produced by eye blinking

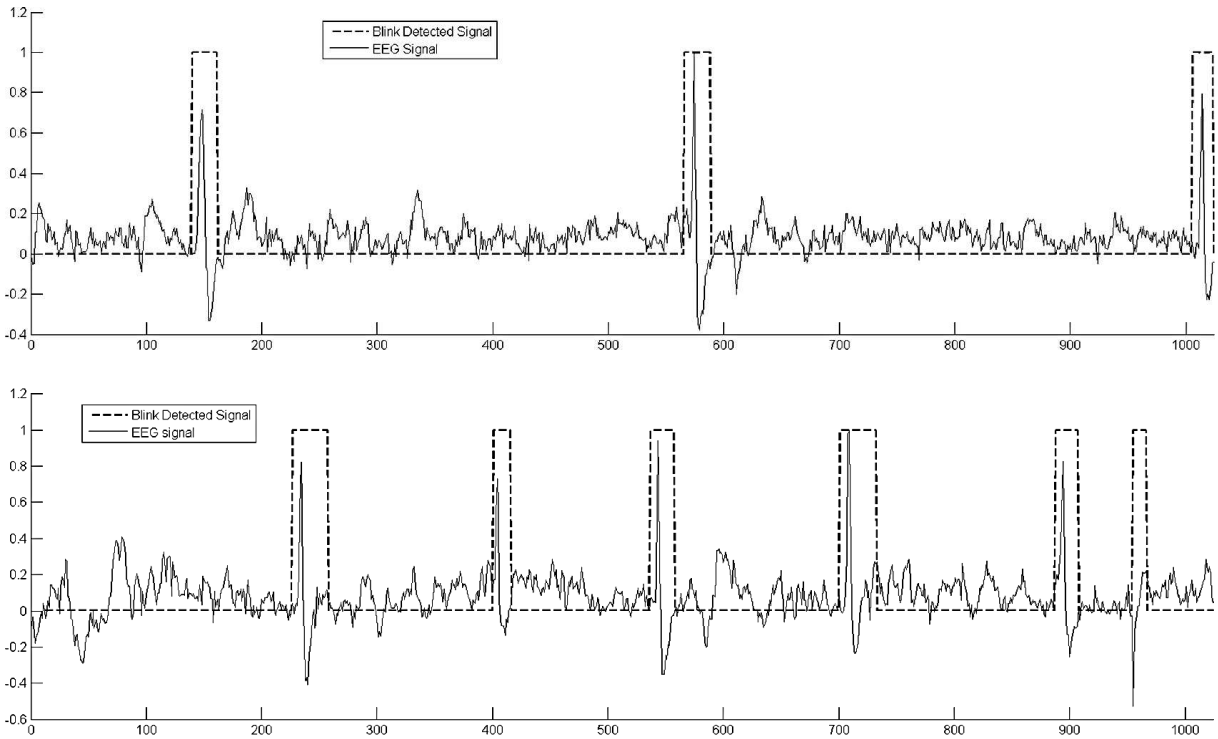


Figure 6: Examples of real-time detection of eye blinking with several analysis windows running over a $1000[ms]$ time interval. a) Top: Three eye blinks; b) Bottom: Six eye blinks

5 Results and discussion

Results from this work have been tested with real-time EEG recordings introducing eye-blinking artifacts with different frequencies and amplitudes. Figure 6 shows two recordings with three and six eye blinking incidences of variable occurrences within 1000[ms] or one second time intervals respectively, showing variable amplitudes and duration. This figure proves that the algorithm is robust enough to filter artifacts produced by other sources that do not share similarity in shape, as well as detecting a target pattern in its many versions independent of the number of occurrences regardless of the extent and the level of symmetry in the study window.

Compared to our previous results [16] we have made some major improvements which allowed robust real-time detection in a given time interval including amplitude and duration of the eye blinking occurrences. A new mother wavelet has been designed according to the recorded patterns of the EEG waves displaying eye blinking and the new wavelet decomposition was achieved with four levels that allowed increasing the number of eye blinking artifacts within a given time analysis window demonstrated in Figures 6(a) and 6(b).

The use of a designed mother wavelet improved the global detection process of the searched pattern detection, using different observation periods for analysis, length of patterns, amplitudes and repetitions. As the relative location was obtained using the "blinkwave" transform wavelet, the thresholding and the sum of coefficients were simple and allowed quick obtaining of precise real time results.

6 Conclusions

The use of wavelet transforms in analyzing non-stationary signals, such as electroencephalograms, shows comparative advantages with respect to a conventional Fourier analysis. The design of a mother wavelet was the main contribution to locate the blinking artifacts in the recorded EEG signals. The paper shows in detail how to design a mother wavelet for a given case. Among the advantages of using wavelet signal processing is the multi-resolution analysis, which involves looking at multiple versions of the blinking patterns. The design of a mother wavelet used in the context of wavelet analysis incorporates features of shape, amplitude, and phase shift for detecting patterns using one-dimensional wavelet transform thus ensuring a robust detection and recognition technique of eye blinking phenomena in EEG signals. This work also contributes to bioengineering with respect to brain computer interfaces or BCI, allowing real-time detection of eye blinking. Two major applications can be visualized: improvement of current BCI devices and awareness detection systems. In addition, these eye blinking patterns, generated voluntarily by the user, can translate into simple commands for helping people with disabilities, e.g., commanding a wheelchair in forward or reverse mode. From the results of our work, we conclude that detection of the eye-blinking artifacts in EEG signals can be very helpful now that the timing information of occurrence is well defined. It should be noted that although there are alternative analyses for eye blinking detection from EEG data, those approaches operate off-line, and our method operates in real time. This approach improves respect previous work, in that there are fewer processing steps now and increased precision regarding the timing of the eye blinking events.

Acknowledgment

All authors like to thank the continuous support of the VRAC of the Technologic Metropolitan University of Chile and VRIDEI of the University of Santiago of Chile.

Author contributions

The authors contributed equally to this work.

Conflict of interest

The authors declare no conflict of interest.

Bibliography

[1] Aeschbach, D.; Borbely, A. (1993). All-night dynamics of the human sleep EEG, *Journal of Sleep Research*, 2, 70-81, 1993.

[2] Bayliss, J. (2001). *Flexible Brain Computer Interface*, PhD thesis, Comp. Science Dept., University of Rochester, 2001.

[3] Binnie, C.; Cooper, R.; Maguire, F.; Osselton, J.; Prior, P.; Tedman, B.(2003). *Clinical Neurophysiology*, Elsevier Academic Press, 2003.

[4] Buzsaki, G. (2006). Rhythms of the Brain; *Oxford University Press, Inc.*, ISBN-13 978-0-19-530106-9, 2006.

[5] Chambayil, B.; Singla, R.; Jha, R. (2010). EEG Eye Blink Classification Using Neural Network, *Proceedings of the World Congress on Engineering*, Vol I. WCE 2010, London, U.K. 2010

[6] Daubechies, I. (1992). *Ten Lectures on Wavelets (CBMS-NSF Regional Conference Series in Applied Mathematics)*, Society for Industrial & Applied Mathematics, U.S. 1992.

[7] Dzitac, I.; Vesselenyi, T.; Tarca, R.C. (2011). Identification of ERD using Fuzzy Inference Systems for Brain-Computer Interface, *International Journal of Computers Communications & Control*, 6(3), 403-417, 2011.

[8] Fisch B.J. (1999). *EEG PRIMER Basic Principles of Digital and Analog EEG. 3rd Edition*, Elsevier Academic Press, 1999.

[9] Gilmore R.L. (1994). American electroencephalographic society guidelines in electroencephalography, evoked potentials, and polysomnography, *Journal of Clinical Neurophysiology*, 11, 147, 1994.

[10] Gloor P. (1969). *Hans Berger on the Electroencephalogram of Man. Amsterdam*, Elsevier Publishing Company, 1969.

[11] Mallat, S. (1989). A theory for multiresolution signal decomposition: the wavelet representation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 11(7), 674-693, 1989.

[12] Mallat, S. (2009). A Wavelet Tour of Signal Processing, The Sparse Way. 3rd Ed., *Library of Congress Cataloging-in-Publication. Data*, Elsevier Inc. Burlington, EEUU, 205-206, 2009.

[13] Neurosky Inc. (2016). BCI: Entertainment's Brain Hacking Tool for Control & Monitoring, *http://neurosky.com/wp-content/uploads/2016/06/Control-vs-Monitor.pdf*

[14] Misiti, M.; Misiti, Y.; Oppenheim, G.; Poggi, J.M. (2007); *Wavelets and Their Applications*, ISTE Ltd., 115-131, 2007.

[15] Polkko J. (2007). *A Method for Detecting Eye Blinks from Single-Channel Biopotential Signal in the Intensive Care*, Unit. Master's Thesis, 2007.

[16] Salinas R., E. Schachter and M. Miranda. (2012). Recognition and Real-Time Detection of Blinking Eyes on Electroencephalographic Signals Using Wavelet Transform, *Lecture Notes in Computer Science*, 7441, 682-690, 2012.

[17] Senthil Kumar, P.; Arumuganathan, R.; Sivakumar, K.; Vimal, C. (2008). Removal of Ocular Artifacts in the EEG through Wavelet Transform without using an EOG Reference Channel, *International Journal of Open Problems in Computer Science and Mathematics*, 1, 188-200, 2008.

[18] Sharbroug, F.; Chatrian, G.E.; Lesser,R.P.; Luders, H.; Nuwer, M.; Picton, T.W. (1991). American Electroencephalographic Society Guidelines for Standard Electrode Position Nomenclature. *Journal of Clinical Neurophysiology*, 8, 200-202, 1991.

INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL
ISSN 1841-9836, e-ISSN 1841-9844, 14(3), 388-400, June 2019.

Manufacturing Process Monitoring in Terms of Energy Management Improving

A. Simo, C. Barbulescu, S. Kilyeni, C. Dragos

Attila Simo*, Constantin Barbulescu, Stefan Kilyeni
Power Systems Department
Politehnica University Timisoara, Romania
attila.simo@upt.ro, constantin.barbulescu@upt.ro, stefan.kilyeni@upt.ro
*Corresponding author: attila.simo@upt.ro

Claudia Dragos
Department of Automation and Applied Informatics
Politehnica University Timisoara, Romania
claudia.dragos@aut.upt.ro

Abstract: Seeking out new technologies and deploying them for competitive advantage is a key priority for any company. In this context, Internet of Things (IoT) promise to change the way we live and work. IoT could help us in overtaking the top global challenges, as energy management, agriculture and food security, environment and natural resources security, etc., but to achieve this vision, "things" need to sense their environment and share this information among them as well as with us to offer intelligent decision-making. Nowadays, IoT become even more interesting due to the existence of Low Power Wide Area (LPWA) networks. LPWA is a key component of a wider IoT technology wave. In this paper, a LoRaWAN based solution is presented, for manufacturing process monitoring, in order to improve energy management in plants. The authors present a minimally invasive device, with which many product manufacturing data can be obtained for analysis and further improvements. This paper is an extended variant of [14][a].
Keywords: intelligent decision-making, LoRaWAN communication protocol, wireless sensor network.

[a]Partially reprinted and extended, with permission based on Licence Number 4587520206486 © IEEE, from "2018 7th International Conference on Computers Communications and Control (ICCCC)."

1 Introduction

This paper presents a part of a more complex work from the Low Power Wide Area (LPWA) networks field, ongoing in the Power Systems Department from Politehnica University Timisoara, in collaboration with ETA2U Innovation.

The current paper is an extended version of [14]. It deals with energy management improvement, based on wireless sensor network and the LoRaWAN communication protocol, dedicated especially to factories with industrial buildings placed over a large geographical area (6 km distance between buildings). In [14] it was presented a battery powered non-invasive device, isolated from any other equipment on the production lines, developed by ETA2U Innovation in collaboration with Politehnica University Timisoara. The job of this device was to collect the data about the manufacturing process, to transmit them over long distances using the LoRaWAN communication protocol. In this paper the new version of this device is presented. This new device brings some improvements, like communication via Ethernet and optocoupler for connections with programmable logic controllers.

LPWA represent a new communication paradigm, which will complement traditional cellular and short-range wireless technologies in addressing diverse requirements of IoT applications.

LPWA technologies offer unique sets of features including wide-area connectivity for low power and low data rate devices, not provided by legacy wireless technologies [1]. These type of networks are unique because they make different tradeoffs than the traditional technologies prevalent in IoT landscape such as short-range wireless networks (Z-Wave, Zig-Bee) [13], legacy wireless local area networks (WLANs), and cellular networks (GSM, LTE), etc. [8].

The legacy non-cellular wireless technologies are not ideal to connect low power devices distributed over large geographical areas because the range of these technologies is limited to a few hundred meters at best. The devices, therefore, cannot be arbitrarily deployed or moved anywhere, which is a requirement for many applications for logistics, smart city and agriculture [17]. Legacy WLANs, on the other hand, are characterized by shorter coverage areas and higher power consumption for machine-type communication. The range of these technologies is extended using a dense deployment of devices and gateways connected using multihop mesh networking [4].

It is obvious, due to the advantages, the LPWA filed is a priority for the mobile industry, and initiatives have been ongoing for several years in an attempt to deliver standards that will enable mobile operators to offer LPWA-like connectivity. With a range of a few to tens of kilometers and battery life of ten years and be-yond, LPWA technologies are promising for the Internet of low-power, low-cost, and low throughput things [9]. A very long range of LPWA technologies enables devices to spread and move over large geographical areas, thus IoT and M2M devices connected by LPWA technologies can be turned on anywhere and anytime to sense and interact with their environment [5].

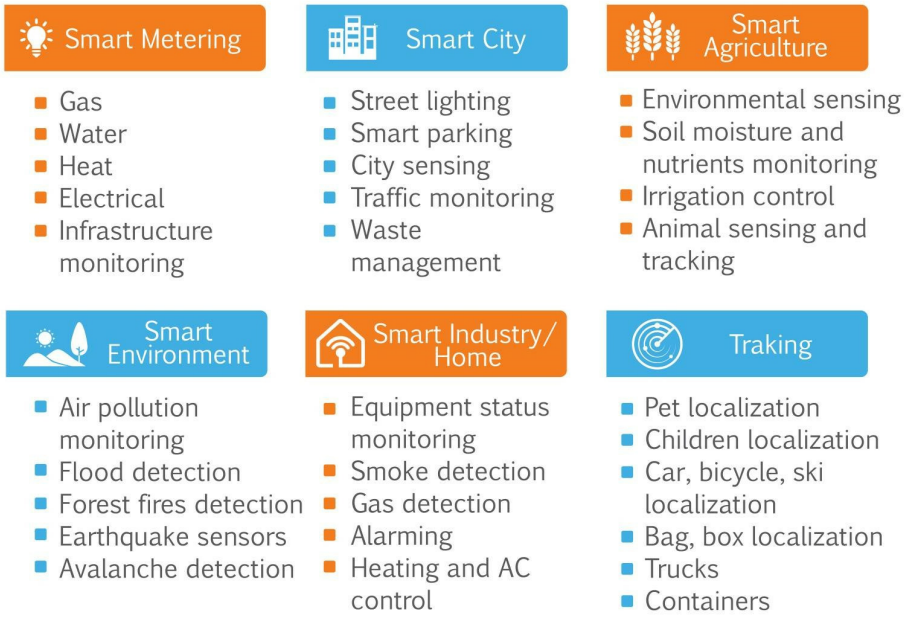


Figure 1: Applications of LPWA technologies across different sectors

Fig. 1 and Fig. 2 highlights the variety of applications across several business sectors that can exploit Low Power Wide Area technologies to connect their end devices. These sectors include smart city, smart grid, smart metering, logistics, industrial monitoring, agriculture, wildlife monitoring, environment monitoring, tracking etc.

It is worth clarifying that LPWA technologies achieve long range and low power operation at

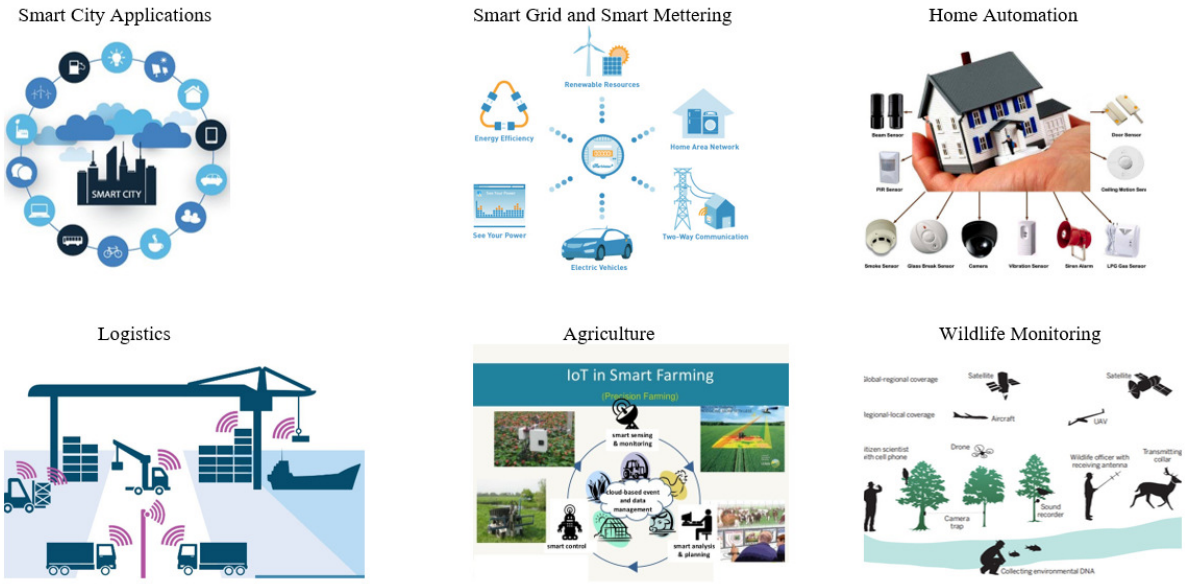


Figure 2: LPWA technologies in different sectors

the expense of low data rate (typically in orders of tens of kilobits per seconds) and higher latency (typically in orders of seconds or minutes). Therefore, it is clear that LPWA technologies are not meant to address each IoT use case. Specifically, LPWA technologies are considered for those use cases that are delay tolerant, do not need high data rates, and typically require low power consumption and low cost. The Internet of Things, as a new growth engine of the information and communications technology industry, has sparked global enthusiasm.

IoT promise to change the way we live and work. IoT could help us in exceeding the top global challenges due to energy crisis, resource depletion, environmental pollution, etc. To achieve this vision, "things" need to sense their environment, share this information among them as well as with us to offer intelligent decision making.

Nowadays IoT become even more interesting due to the existence of LPWA networks.

Different studies have forecasted a growth in volume and revenue of IoT in the next several years. Number of connected devices and consumer electronics will surpass the number of human subscribers using mobile phones, personal computers, laptops and tablets by 2020 [13]. Moving forward, by 2024, the overall IoT industry is expected to generate 4.3 trillion dollars' revenue across different sectors such as device manufacturing, connectivity, and other value added services [15]. Recent improvements in cheap sensor and actuation technologies along with an emergence of novel communication technologies are all positive indicators, supporting the forecasted trends [11].

Nowadays, energy efficiency in manufacturing processes provides several advantages to companies, such as cost saving notwithstanding rising and volatile energy prices. In [6] the authors define green products as those that have been manufactured while consuming as little energy as possible – not just products which consume less energy when used by the customer. Consequently, the best practices that aim to reduce energy consumption during manufacturing process are increasingly important to today's manufacturing companies.

In many companies, energy management practices at manufacturing process levels suffer due to lack of awareness of energy consumption behavior. In fact, energy savings are expected to be achievable both from improvements in energy efficiency of specific manufacturing processes, as well as from the usage of innovative energy monitoring systems and management approaches [16].

New emerging autonomous technologies, such as IoT, are enhancing monitoring of manufacturing processes. IoT technology (e.g. smart meters, sensors) provide awareness of energy consumption patterns by collecting real-time energy consumption data.

The availability of real-time energy consumption data gives us several opportunities to reduce energy consumption by enabling and enhancing energy-efficient practices in manufacturing process management [15].

The focus of this research is to build a more flexible device, as the one presented in [14], that helps us track manufacturing process in order to increase the efficiency of electricity consumption - e.g. identifying causes and reducing the downtimes between producing two elements.

In the 1st section generalities about LPWA are discussed, the 2nd one refers to the LoRa technology and the LoRaWAN communication protocol. In 3rd section the new monitoring device is presented, and the results and discussions are tackled. Finally, conclusions are synthesized within the 4th section.

2 The LoRa technology

LoRa is a proprietary physical layer for LPWA connectivity. A physical layer technology that modulates the signals in SUB-GHz ISM band using a proprietary spread spectrum technique developed and commercialized by Semtech Corporation. A bidirectional communication is provided by a special chip spread spectrum technique, which spreads a narrow band input signal over a wider channel bandwidth. The resulting signal has noise like properties, making it harder to detect or jam.

A special interest group constituted by several commercial and industrial participants dubbed as $LoRa^TM$ Alliance proposed LoRaWAN, an open standard defining architecture and layers above the LoRa physical layer. A simple ALOHA scheme is used at the MAC layer that in combination with LoRa physical layer enables multiple devices to communicate at the same time but using different channels and/or orthogonal codes (spreading factors). End devices can hop on to any base station without extra signaling overhead. The base stations connect end devices via a backhaul to network server, the central unit of the LaRaWAN system that suppresses duplicate receptions, adapts radio access links, and forwards data to suitable application servers (fig. 3). Application servers then process the received data and perform user defined tasks.

Figure 3: Architecture of a LoRa based application

LoRaWAN anticipates that the devices will have different capabilities as per application

requirements. Therefore, LoRaWAN defines three different classes of end-devices, all of which support bidirectional communication but with different downlink latency and power requirements:

- Class A device achieves the longest lifetime but with the highest latency. It listens for a downlink communication only shortly after its uplink transmission.
- Class B device, in addition, can schedule downlink receptions from base station at certain time intervals. Thus, only at these agreed-on epochs, applications can send control messages to the end devices (for possibly performing an actuation function).
- Class C device is typically mains-powered, having capability to continuously listen and receive downlink transmissions with the shortest possible latency at any time.

LoRaWAN standard uses symmetric-key cryptography to authenticate end devices with the network and preserve the privacy of application data [13].

Several studies evaluated LoRaWAN in real world environments including outdoor and even indoor settings. For example, a study in [9] observed 15 km and 30 km communication ranges for LoRaWAN on ground and water respectively in Oulu Finland.

In another study [11], end devices transmitted at 14 dBm using highest spreading factor to the base station that was located within 420 m radius. The packet delivery ratio at the base station is recorded to be 96.7%.

In [10] the authors evaluated the performance of LoRaWAN unconfirmed uplink data frames in an indoor environment. They first showed the limitations in term of periodicity and size of data because of the ISM band regulation in a default channel configuration. Such regulation also limits the maximum amount of data that can be sent per day. Then, they evaluated the signal quality received from different locations, in order to verify the feasibility to cover an entire building with the LoRaWAN technology. The difference in the composition of walls between the rooms and the lab floors had not much impact on the quality of the transmission and packet loss. Only communications with the basement experienced degradations. They also showed that the data rate can be a factor of loss and should be selected appropriately when configuring end-device. Finally, they showed the average current consumption of a LoRa mote and how the used data rate can affect the global energy consumption. [7] propose an architecture of smart health monitoring system for smart factory and proposes the methodology for maximizing the productivity.

In [2], first part of the paper, the authors are dealing with electrical and heat energy multipoint measurements and consumption reduction options. They highlights current studies, obtained findings and first results, focusing on industrial and office type building electrical energy consumption, definition of problems and description of possible experimental solution. In [3], the second part, they present the first measurement data and analysis results and prove the need for a mathematical model and prediction algorithm development for improved and energy efficient building regulation processes. They conclude that LoRAWAN network is well suited for the non-dynamic sensor data transfer, but not for detailed electrical power consumption, thus a new approach for data aggregation is needed, by means of embedded algorithms already within the power measurement equipment.

The use of a long-range (LoRa) technology, originally developed for IoT, is investigated with the aim of implementing DMSs is presented in [12]. After the conclusion that LoRa and LoRaWAN architectures show a good match with measurement systems, this paper focuses on the characterization of time-related performance indicators that are important for distributed systems. The experimental results show the capability of low-cost transceiver to schedule the

transmission of frames with a standard uncertainty less than 3 μs and an acceptable long-term clock stability (Allan Deviation) of commercial available devices (nodes and packet forwarders) for application such as smart metering, smart building, and process industry.

3 Results and discussions

In Romania, unfortunately in case of many plants, the industrial buildings are placed over large geographical area, few kilometers, in most cases these are not equipped with an Internet connection or other related facilities (due to the impact of the investments). Monitoring the manufacturing process using XXI century technologies is hampered by these disadvantages.

Decision-makers, process engineers are always looking to improve the manufacturing process. They want to have clear production statistics - how many equipment were produced on the line per hour, per shift handover, per day, seeking the human or machines efficiency. They are also interested about the downtime between the production of two equipment, in term of energy management. In several cases, as we received requests decision-makers want these data to be transmitted over long distances. In the absence of the Internet, these things are difficult to achieve.

The main purpose of [14] is to provide solution for some particular requirements from such plants. They need a monitoring device to increase energy management efficiency, which at the same time is able to transmit data over long distances, over the air.

In this direction, in the laboratories of ETA2U - IoT solutions integrator company – by ETA2U Innovation was developed and tested a monitoring device to meet the expectations of these factories.

Minimal intervention on production lines, low consumption of our equipment and long-range data transmission were the special requirements of customers [14].

The developed equipment consists of 3 main parts:

- Non-invasive Split Core Current Transformer, type YHDC 100A-STC-013-000 – which takes the parameters from one phase of the circuit.
- Arduino Mega (ATmega 1280 - Fig. 4) with a Dragino Lora Shield v1.4 (Fig. 5) – for reading, processing and transferring the collected data using LPWA technology.
- Own developed Printed Circuit Board (PCB) (fig. 13) – through which the connection to the technological process is made. Mainly this part of the device has been further developed. An Ethernet port and optocouplers have been added, the last ones for connection to the manufacturing process via programmable logic controllers.e.

This type of microcontroller has been chosen due to the energy efficiency and cost benefits. Also, it PowerSaving, SleepMode, SleepModeWatchDog features have played an important role in our decisions. It is user friendly and is easy to implement applications on it. Thanks to these functions, this microcontroller reaches a typical consumption of about 0.013 mA, by which we partly responded to customer requirements.

One of the customer's special requirement it was that after the data has been collected, to be transmitted over large geographical area. We did not have the possibility of transmission the data via Internet. For this case, we have chosen to use LoRa technology (again taking into account that we need to have low energy consumption). We chose to add to our device the Dragino LoRa Schield v1.4. This device is a long-range Arduino transmitter emitter that works with an open source library. It allows the user to send and receive data at long distances with low transfer rate.

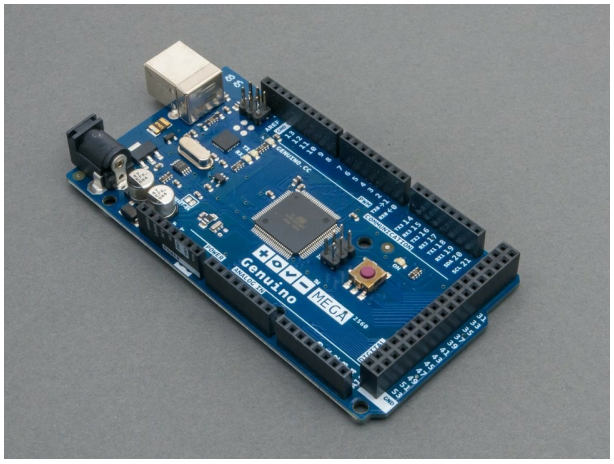


Figure 4: Arduino Mega - for reading, processing and transfer data

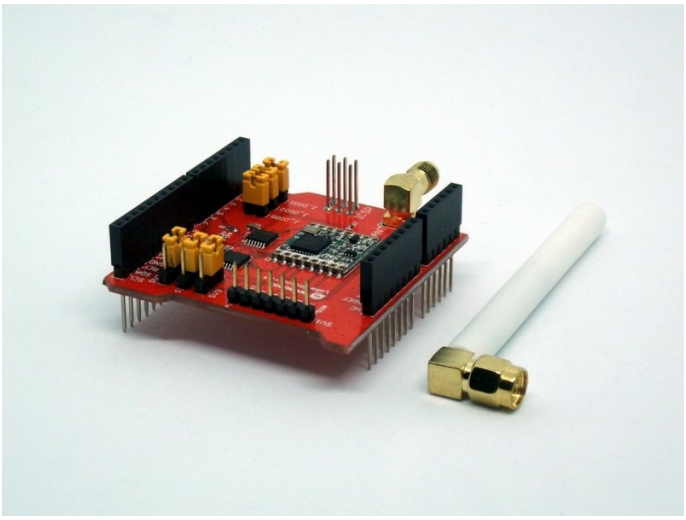


Figure 5: Dragino Lora Shield - long range data transmission

This shield (fig. 5) is based on a RFM98W radio frequency transmitter, which can be used with professional wireless sensor systems applications. Thanks to its high sensitivity up to –148dBm combined with a + 20dBm power amplifier, it makes it the best choice for remote communication applications.

The PCB between the current transformer and the microcontroller has a double role:

- The first is to "wake up" the microcontroller – because when there is no activity on the production line the controller enters in standby mode. For this, we used an LM393 comparator to generate an interruption on Arduino Mega's number 21 digital pin. This is done by the voltage difference between the reference + and fixed - terminals. When a consumer voltage value is connected to the terminal – it will exceed the reference value at terminal +. This will generate an output signal that will "wake up" the microcontroller.
- The second one is to connect the current transformer sensor to the microcontroller. The output signal from the current transformer needs to be conditioned so it meets the input requirements of the controller analog inputs.

Fig. 6 shows the scheme of this intermediate circuit [14], which has the following components

(each of them calculated, according to our needs):

- 3 x 10 kOhm resistor;
- 1 x 97 Ohm resistor (burden resistance to close the circuit);
- 1 x 22 kOhm resistor;
- 1 x 50 kOhm resistor;
- 1 x 10 μF capacitor;
- 1 x 1000 μF capacitor;
- 1 x 27 μF capacitor;
- 1 x LM393 comparator.

The circuit consists of five main parts:

- the current transformer sensor and the burden resistor;
- the comparator;
- the biasing voltage divider;
- Ethernet port;
- optocouplers.

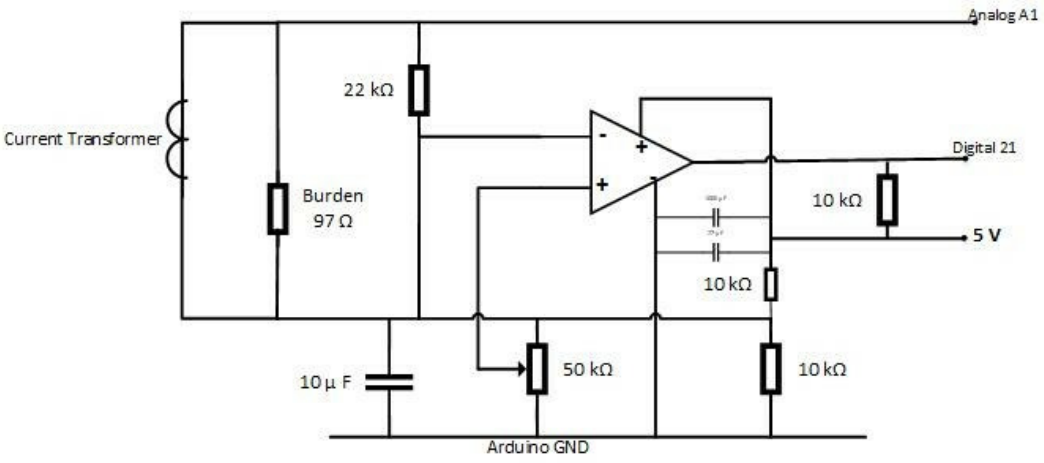


Figure 6: "Wake up" and measuring circuit scheme

After calculating the parameters, we built the prototype of the device (Fig. 7, Fig. 8).

Until we reached the final parameter values we had some obstacles to pass in the testing and validation phase. Due to the noise on the initial circuit, "waking up" was done also without connecting a consumer (Fig. 9). To reduce the noise, on the comparator feed we put 2 capacitors, one electrolytic of 1000 μF (Fig. 10) and one ceramic of 27 μF (Fig. 11) . Fig. 12 shows the circuit noise with both capacitors.

Another problem was that sometimes the signal was not strong enough to generate the necessary interruption. To solve this problem, we put a 10 kOhm resistance between the output

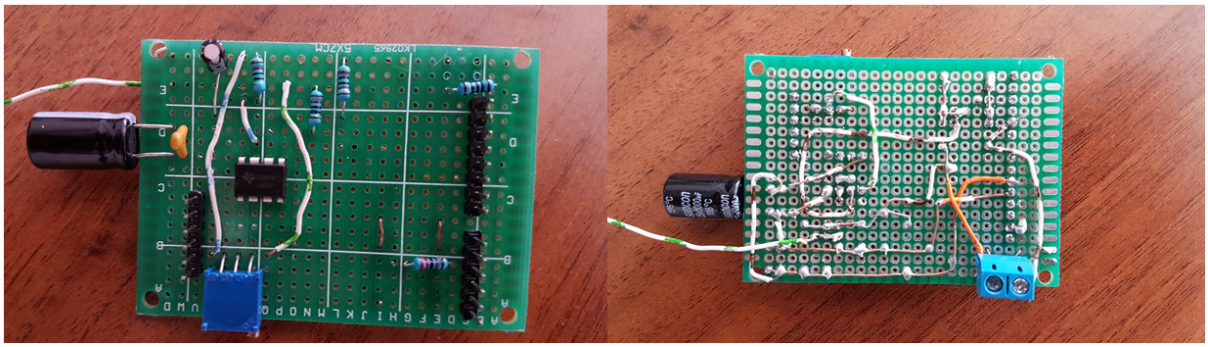


Figure 7: Circuit prototype

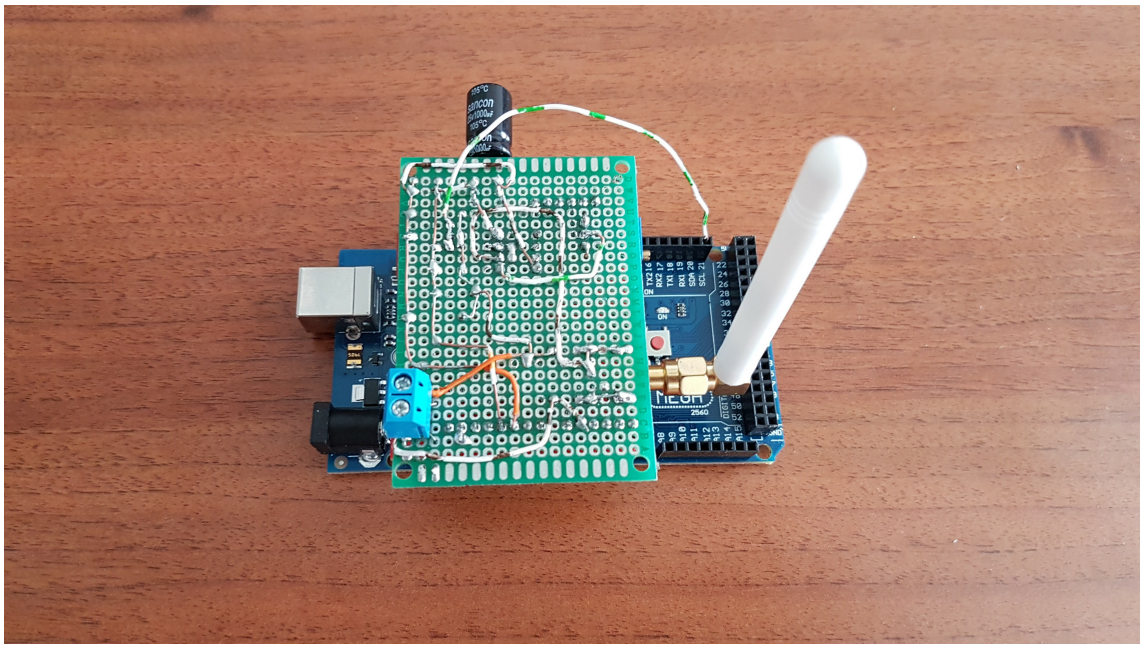


Figure 8: Device prototype

of the amplifier and the 5 V supply of the microcontroller. Fig. 13 shows the final shape of the circuit.

Considering these corrections, the final form of this device (Fig. 14) was mounted on production lines in several factories, collecting and sending data through LoRaWAN to a gateway behind which is an application that allows data analysis, graphs, helping the process engineers optimizing the process and the energy consumption.

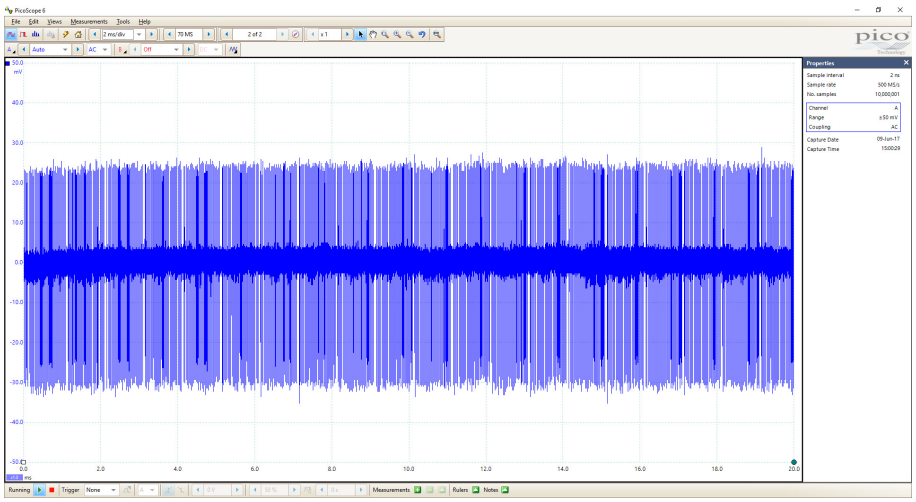


Figure 9: Circuit noise without capacitors (2 ms/div)

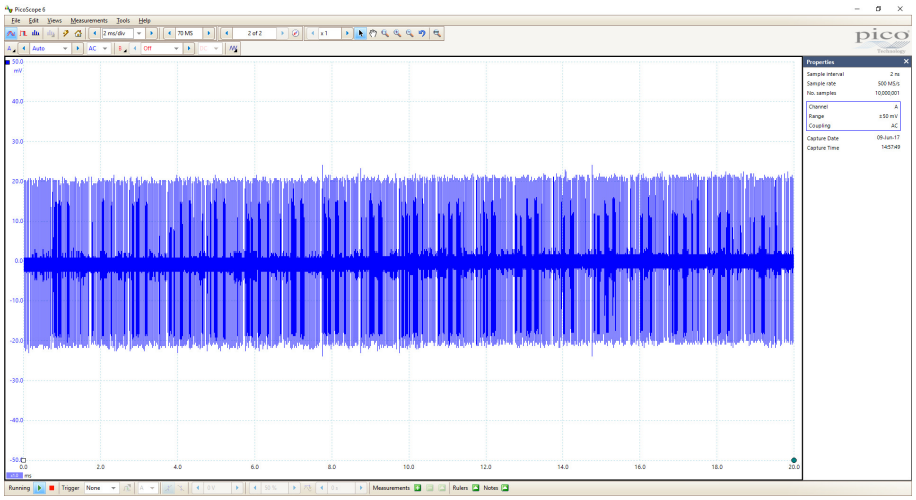


Figure 10: Circuit noise with 1000 μF electrolytic capacitor (2 ms/div)

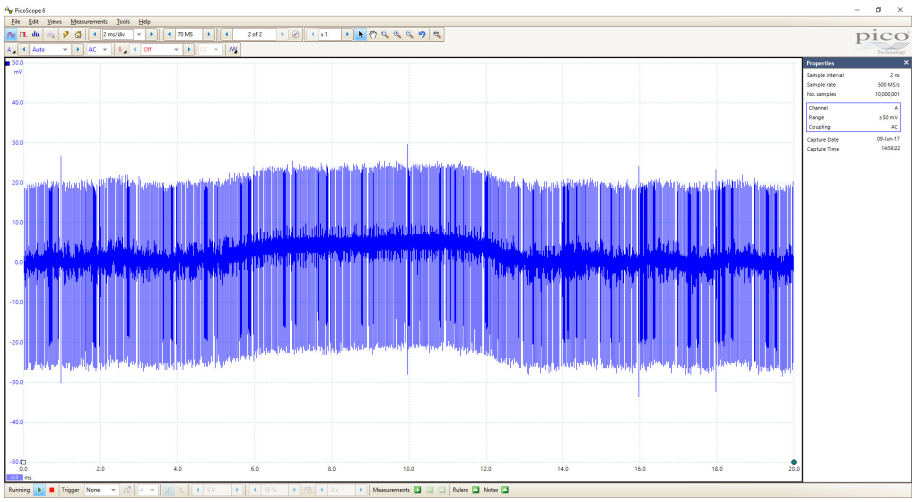


Figure 11: Circuit noise with 27 μF ceramic capacitor (2 ms/div)

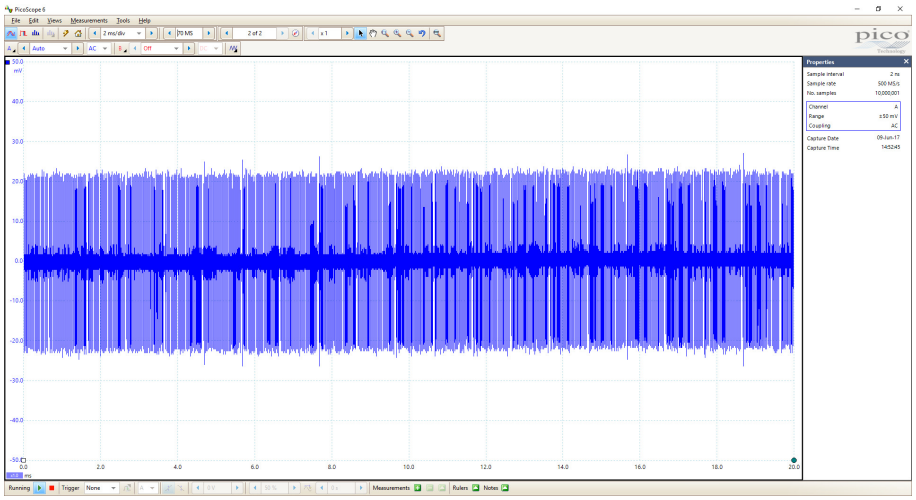


Figure 12: Circuit noise with both capacitors (2 ms/div)

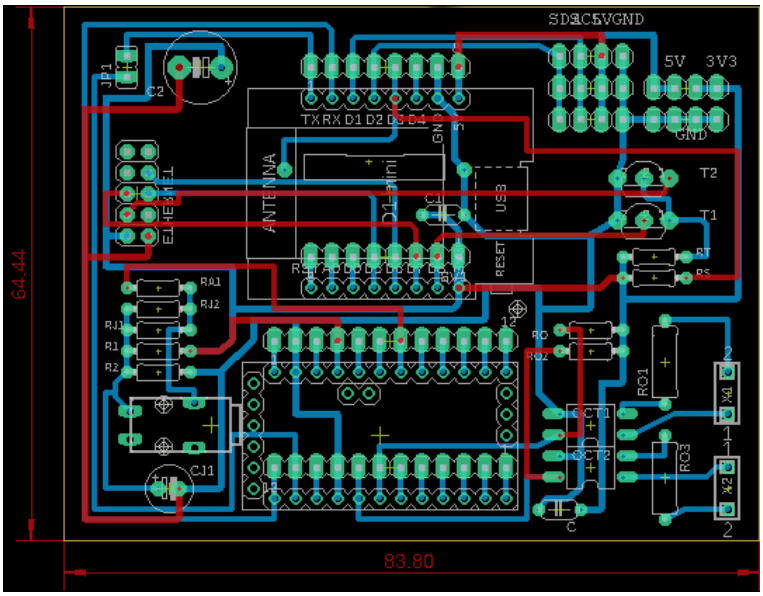


Figure 13: The final form of the circuit

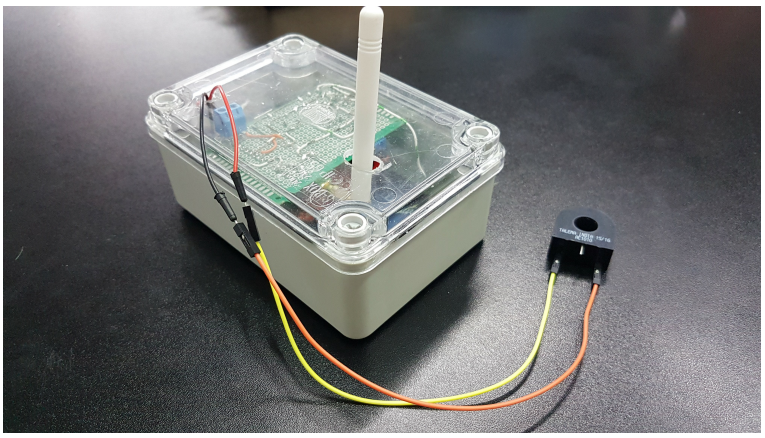


Figure 14: The final form of the device

4 Conclusions

This paper addresses a topic that companies need to consider once they plan to improve the energy efficiency of their manufacturing facilities and achieve this goal thought IoT solutions. These solutions enable a very high level of awareness, being capable of being flexibly installed and collecting large quantities of energy-related data (and beyond), almost in real-time. For this reason, it is highly important to design in advance how such IoT energy monitoring solutions have to be included in the company's energy management approach.

The developed device offers a novel perspective on integrating "campus"/isolated factories (not only) in future energy management trend, using IoT solutions. Moreover, the new version of the device allows more than the previous one, such as Ethernet connection or the connection to the technological process via the programmable logic controllers, using optocouplers.

There are several situations, as is the case with these companies when more is needed, not only sensors connected via Internet; wide area coverage, low power consumption, and inexpensive wireless connectivity are necessary to achieve the goal.

The energy management practices and the proposed device offer a novel perspective on integrating energy data into manufacturing process management and related decisions at the operational level within these companies.

Funding

This work was partial supported by the Politehnica University Timisoara research grants PCD-TC-2017. The authors are addressing special thanks to the members of ETA2U Innovation.

Author contributions

The authors contributed equally to this work.

Conflict of interest

The authors declare no conflict of interest.

Bibliography

[1] Augustin, A.; Yi, J.; Clausen, T.; Townsley, W.M. (2016). A Study of LoRa: Long Range & Low Power Networks for the Internet of Things, *Journal of Sensors*, 16(9), 1–18, 2016.

[2] Avotins, A.; Senfelds, A.; Apse-Apsitis P.; Podgornovs, A. (2017). IoT solution approach for energy consumption reduction in buildings: Part 1. Existing situation and problems regarding electrical consumption, *Proceeding of the 58th International Scientific Conference on Power and Electrical Engineering of Riga Technical University (RTUCON)*, 31–36, 2017.

[3] Avotins, A.; Podgornovs, A.; Senfelds, A.; Vegeris M. (2018). IoT Solution Approach for Energy Consumption Reduction in Buildings: part 2. Measurement Setup and Practical Data Analysis, *Proceedings of the17th International Scientific Conference on Engineering for Rural Developmen*, 923–929, 2018.

[4] Berthelsen, E.; Morrish, J. (2015). *Forecasting the internet of things revenue opportunity*, Tech. Rep., 2015.

[5] Cattani, M.; Boano, C.A.; Romer, K. (2017). An Experimental Evaluation of the Reliability of LoRa Long-Range Low-Power Wireless Communication, *Journal of Sensor and Actuator Networks*, 6(2), 1–18, 2017.

[6] Garetti, M.; Taisch, M. (2012). Sustainable manufacturing: trends and research challenges, *Journal ofProdustion Planning & Control*, 23(2), 83–104, 2012.]

[7] Kharel, J.; Shin, S. Y. (2017). Smart Health Monitoring System of Employee for Smart Factory, *Proceedings of Symposium of the Korean Institute of communications and Information Sciences*, 1–6, 2017.

[8] Nokia (2016). *LTE evolution for IoT connectivity*, Tech. Rep., 2016.

[9] Petajajarvi, J.; Mikhaylov, K.; Roivainen, A.; Hanninen, T.; Pettissalo M. (2015). On the coverage of LPWANs: range evaluation and channel attenuation model for lora technology, *Proceedings of the 14th International Conference on ITS Telecommunications (ITST)*, 55–59, 2015.

[10] Neumann, P.; Montavont, J.; NoĂŤl, T. (2016). Indoor Deployment of Low-Power Wide Area Networks (LPWAN): a LoRaWAN case study, *Proceedings of the 2th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*, 54–58, 2016.

[11] Petjjrvi, J.; Mikhaylov, K.; Hmlinen, M.; Iinatti, J. (2016). Evaluation of LoRa LPWAN technology for remote health and wellbeing monitoring, *Proceedings of the 10th International Symposium on Medical Information and Communication Technology (ISMICT)*, 1–5, 2016.

[12] Rizzi, M.; Ferrari, P.; Flammini, A.; Sisinni, E. (2017). Evaluation of the IoT LoRaWAN Solution for Distributed Measurement Applications, *Journal of Instrumentation and Measurement*, 66(12), 1–18, 2017.

[13] Raza, U.; Kulkarni, P.; Sooriyabandara M. (2017). Low Power Wide Area Networks: An Overview, *Journal of Communications Surveys & Tutorials*, 19(2), 855–873, 2017.

[14] Simo, A.; Barbulescu C.; Kilyeni, S.; Dragos, C. (2018). LoRa based Energy Efficiency Improving in Manufacturing Processesn; *Proceeding of the 7th International Conference on Computers Communications and Control (ICCCC)*, 156–161, 2018.

[15] Shrouf, F.; Miragliotta, G. (2015). Energy management based on Internet of Things: practices and framework for adoption in production management, *Journal of Cleaner Production*, 100, 235–246, 2015.

[16] Weinert, N.; Chiotellis, S.; Seliger G. (2011). Methodology for planning and operating energy-efficient production systems, *CIRP Annals*, 60(1), 41–44, 2011.

[17] Xiong, X.; Zheng, K.; Xu, R.; Xiang, W.; Chatzimisios, P. (2015). Low power wide area machine-to-machine networks: key techniques and prototype, *IEEE Communications Magazine*, 53(9), 64–71, 2015.

INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL
ISSN 1841-9836, e-ISSN 1841-9844, 14(3), 401-418, June 2019.

Enhanced Fireworks Algorithm-Auto Disturbance Rejection Control Algorithm for Robot Fish Path Tracking

X. Song, S. Gao, C. Chen, Z. Gao

Xiaoru Song
School of Electronic and Information Engineering, Xian Technological University
No.2 of Xuefu Middle Road, Xian, Shaanxi, China
masha0422@163.com

Song Gao*
School of Electronic and Information Engineering, Xian Technological University
No.2 of Xuefu Middle Road, Xian, Shaanxi, China
*Corresponding author: gao2223@qq.com

Chaobo Chen
School of Electronic and Information Engineering, Xian Technological University
No.2 of Xuefu Middle Road, Xian, Shaanxi, China
17777108@qq.com

Zepeng Gao
School of Electronic and Information Engineering, Xian Technological University
No.2 of Xuefu Middle Road, Xian, Shaanxi, China
1067456190@qq.com

Abstract: The robot fish is affected by many unknown internal and external interference factors when it performs path tracking in unknown waters. It was proposed that a path tracking method based on the EFWA-ADRC (enhanced fireworks algorithm-auto disturbance rejection control) to obtain high-quality tracking effect. ADRC has strong adaptability and robustness. It is an effective method to solve the control problems of nonlinearity, uncertainty, strong interference, strong coupling and large time lag. For the optimization of parameters in ADRC, the enhanced fireworks algorithm (EFWA) is used for online adjustment. It is to improve the anti-interference of the robot fish in the path tracking process. The multi-joint bionic robot fish was taken as the research object in the paper. It was established a path tracking error model in the Serret-Frenet coordinate system combining the mathematical model of robotic fish. It was focused on the forward speed and steering speed control rate. It was constructed that the EFWA-ADRC based path tracking system. Finally, the simulation and experimental results show that the control method based on EFWA-ADRC and conventional ADRC makes the robotic fish track the given path at $2.8s$ and $3.3s$ respectively, and the tracking error is kept within plus or minus $0.09m$ and $0.1m$ respectively. The new control method tracking steady-state error was reduces by 10% compared with the conventional ADRC. It was proved that the proposed EFWA-ADRC controller has better control effect on the controlled system, which is subject to strong interference.
Keywords: Active disturbance rejection control, robot fish, path tracking, enhanced fireworks algorithm.

1 Introduction

The use of bionic robot fish is deepened in narrow or dangerous underwater environments, such as monitoring, underwater archaeology, underwater equipment overhaul, military investigation, etc. The new challenges are imposed on the intelligent requirements of robot fish. The path tracking is involved the kinematics and complex hydrodynamic characteristics of the bionic robot

fish body in the unknown waters, as well as coupling, system nonlinearity, and the influence of various unknown internal and external disturbance factors. Many domestic and overseas scholars have conducted a lot of research on the tracking of robotic fish, and have achieved certain results. The path tracking only requires the robot fish to converge to the desired path. The desired position is not constrained by the time condition. It can make the robot fish complete the path tracking task at the desired speed, so it can smoothly converge to the desired path, and the control input saturation is not easy to occur. Therefore, it is beneficial to improve the endurance of the robot fish. However, the design of the controller becomes more difficult due to external disturbances, limited lateral and vertical motion limitations, high nonlinearity and uncertainty of hydrodynamic coefficients. The literature [7, 25] establishes the horizontal plane tracking error equation based on the virtual wizard, and then designs the forward speed controller based on Lyapunov theory and backstepping method to realize the path tracking of the horizontal plane. In [26], the tracking error equation is established by introducing the approach angle and virtual guide, and then the horizontal path tracking controller is designed based on Lyapunov theory and backstepping method. Phi Luan proposed a robotic fish dynamics model with a non-uniform flexible tail (NFT). It was studied the relationship between input torque and speed. The dynamic model of the robotic fish was verified through experiments. But the problem has not been resolved the motion control problem of the robotic fish [1]. A point-and-point stabilization algorithm was proposed for robotic fish. This strategy aims to eliminate the direction error and distance error between the starting point and the target position of the robotic fish. It is not considered the real-time position control of the robotic fish during the tracking process. In addition, the model is not accurate enough and complicated water wave interference. The actual control effect is not ideal. A fuzzy PID control method was proposed. It was optimized the mathematical model of the robot fish. The robot deep-motion tracking was completed by the fuzzy self-tuning method. The robotic fish has good anti-interference during the tracking process. And it does not require an accurate mathematical model of the robotic fish. It can be effectively implemented by the human experience and control strategies. However, the fuzzy control setting rules require expert knowledge or a large number of experiments to obtain. It cannot be widely applied to the control of the robot fish [9]. A sliding-mode path tracking was proposed [11]. It is well adapted to the nonlinear, indeterminate state of the model. The simulation results show it has better control accuracy. However, it is prone to chattering effect due to inertia and time lag factors. A new method was proposed a curve path tracking based on line-of-sight navigation by Signe Moe considering the interference of the ocean current environment. The line guidance law and the sliding mode adaptive feedback linearization controller were used for path tracking to obtain good control effects. However, the chattering problem has not been solved [4]. The ADRC control method was proposed for the fish-like finfish. The simulation shows that the method has good dynamic and static characteristics for the fish heading control [7]. It was proposed a path tracking based on ADRC under-actuated AUV. The simulation results show that the chattering phenomenon and overshoot caused by interference is well suppressed [6]. And the ADRC parameters are set according to experience in the paper. It may not be optimal [14, 15].

There are many ADRC parameter tuning methods. The bandwidth-based linearized ADRC was proposed by Gao Z. The good effect of this method is obtained in practical applications. However, this method does not give full play to the advantages of nonlinear mechanism [25]. The steady-state performance and rapidity of the control system are considered comprehensively based on the artificial intelligence-based parameter tuning method. The artificial immune algorithm and particle swarm optimization algorithm were used to adjust the ADRC parameters [5]. It has the disadvantages of time-consuming and inevitable local minimum. The chaotic particle swarm optimization algorithm was to adjust the parameters of the controller [8]. But large time consuming is still exists. The reinforcement learning was proposed to adaptive the ADRC algo-

rithm parameters. It has a long online learning time. the continuous state space is not easy to be processed. When the state is large, the dimension explosion problem is easy to appear [19]. The fireworks algorithm was proposed by Chinese scholar Tan Ying. The algorithm has good performance and high efficiency in solving complex optimization problems [18].

In summary, the enhanced fireworks algorithm is adopted to optimize the parameters of ADRC. And a path tracking algorithm based on EFWA-ADRC is proposed to make the high-quality tracking effect of the robot fish in the path tracking process. The multi-joint bionic robot fish is used as the research object in the paper. The path tracking error equation is established in the SF coordinate system. The expected angle guidance algorithm is introduced to design the required heading angle and the guidance rate of the forward speed control to eliminate the tracking error. It is proved the stability of the error dynamics equation. On this basis, the EFWA-ADRC controller is used to control the forward speed and yaw angle of the robot fish respectively. So the robot fish can track the given path at the desired speed comparing with the conventional ADRC control method to verify the feasibility and reliability of the proposed controller [16,17,20]. The rest of the paper is organized as follows. The problem description is presented in the section 1. The part 2 is the key. It is designed of EFWA-ADRC second-order auto-disturbance path tracking controller. The simulation and experiment of the tracking controller of the EFWA-ADRC algorithm are in the section 3. At last, conclusions are drawn [21,22].

2 Problem description

2.1 Multi-joint robotic fish mathematical model

The multi-joint bionic robot is a special underwater robot. It mimics the way in which hundreds of millions of years of fish have evolved. The propulsion model of the robotic fish is proposed by the literature [2]. The four-joint robotic fish is used as the research object in the paper. It consists of one body link, two movable tail links and a passive caudal fin link. That is four links to approximate the state of fish movement at each moment. Fig.1 is a four-joint robotic fish structure. The position of the robot fish each link is obtained by digital fitting method [13].

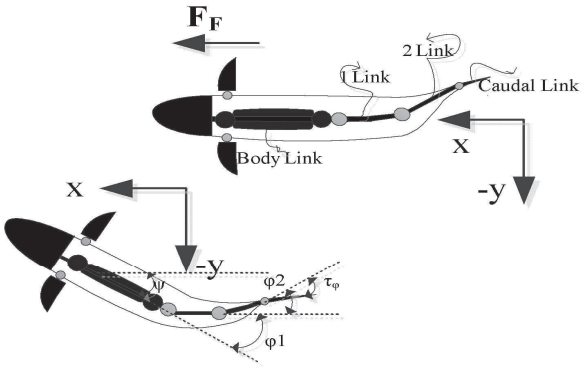


Figure 1: The structure diagram of the robot fish

There are strong coupling between the connecting rods of the robot fish in practical applications. It forces the robot fish analysis is very complicated due to the limitation of the actuator and mechanical structure. The thrust most generated by the caudal fin contributes during the movement of the robotic fish.

In order to simplify the research, the two joints force analysis of the tail is mainly carried out, and the dynamic equation of the robot fish is obtained as follows (1).

$$\begin{cases} M_{RH}\dot{v} + C_{RB}(v)\, v = \tau_H \\ \tau_H = -C_A(v_r)\, v_r - D(v_r)\, v - g(\Theta) + A_c(v_c)\, v_r + \tau_C \end{cases} \tag{1}$$

where, τ_H is the thrust of the machine fish to get forward. V_r is the speed of the robot fish relative to the water flow. V_c is the speed of the water flow. $C_{RB}(v)$ is the fish body coefficient, $C_A(v_r)$ is the water flow force matrix.$D(v_r)$ is the total water flow damping matrix. $g(\Theta)$ is a vector matrix containing gravity and buoyancy. $A_c(v_c)$ is the ocean current velocity Gaussian Markov process thrust, and M_{RH} is the total mass of the robotic fish.$M_{RH} = M_{RB} + M_A$, M_R is the mass matrix of the rigid body, and M_A is the additional mass matrix of the fish body. The fine derivation process can be referred to the literature [10].

2.2 Robotic path tracking error model in carrier coordinate system

The robotic fish movement is more complicated. In order to simplify the research, it is assumed that the robotic fish only moves in the horizontal plane, and ignores the robotic yaw, roll, and pitch motion. The following three coordinate systems are used, as shown in Fig.2 in the paper. They are the inertial coordinate system$E-\xi\eta\{I\}$ of the robot fish, the fish body coordinate system $O-xy\{B\}$, and the SF coordinate system $P-x_py_p\{SF\}$. Where, Point$P$ on the SF coordinate system is a free reference point on the reference path, expressed as a virtual target robot. $P-x_{p(s)}$ is the tangent direction of the reference path. $P-y_{p(s)}$ is the normal direction of the reference path . The speed vector of the robot fish is V_t, the speed of each component in the coordinate system of the fish body is expressed as (u,v,r), where (u,v,r) is forward speed. lateral speed, steering speed respectively, where ψ is the angle between $O-x$ and the positive half of the $E-\xi$. Assume that the reference point P (virtual robot) moves along the reference path at speed U_P. The angle between the $P-x_{p(s)}$ axis and the $E-\xi\eta\{I\}$ axis of the ground coordinate system is ψ_p.

The kinematic equation of the robot fish is as follows formula (2)

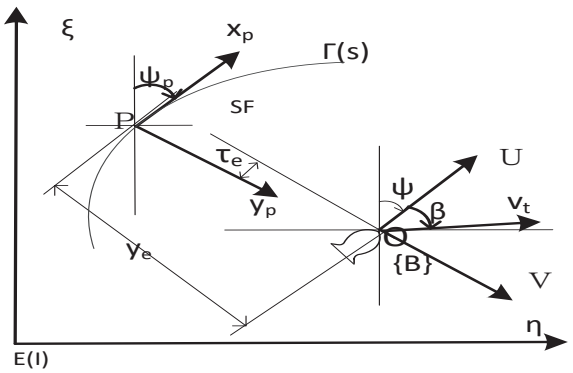


Figure 2: The schematic diagram of robot fish path tracking

$$\begin{cases} \dot{\xi} = cos\psi \cdot u - sin\psi \cdot v \\ \dot{\eta} = sin\psi \cdot u + cos\psi \cdot v \\ \dot{\psi}_{\mathrm{B}} = \mathrm{r} + \dot{\beta} \end{cases} \tag{2}$$

In the formula (2)ψ_B is the angle between the synthetic fish speedV_t and the $E-\xi$. β is the drift angle of the robot fish. It is $\psi = \psi_B + \beta$. To simplify the study, let $\beta = \arctan(v/u)$.

The reference path is described by curve $\Gamma(s)$. The reference point $P\left(x_{p(s)}, y_{p(s)}\right)$ on the curve is uniquely determined by the $\Gamma(s)$ curve. In order to calculate the error angle, the distance from the centroid point O of the robot fish to the P point of the virtual robot is projected to the SF coordinate system. The error dynamics system (τ_e, y_e, n_e) under the coordinate system is obtained, that is the equation (3) is the path following error dynamics equation.

$$\begin{cases} \dot{\tau}_e = \dot{s} - y_e w_p - v_t \cos\psi_e \\ \dot{y}_e = \tau_e w_p - v_t \sin\psi_e \\ \dot{\psi}_e = w_p - r - \dot{\beta} \end{cases} \tag{3}$$

3 Design of EFWA-ADRC second-order auto-disturbance path tracking controller

3.1 Kinematic control rate design

Forward speed is used to control the actual speed V_t. the goal is the forward speed u is maintained at a constant value u_d. That is constant greater than zero throughout the tracking process.

Define the tracking speed error as:

$$u_e = u - u_{\mathrm{d}} \tag{4}$$

The elimination of the tangential error τ_e needs to be controlled by selecting an appropriate U_p such that the tangential error τ_e approaches zero. For the normal error, it needs to be controlled by the angle ψ. The method of the desired angle is employed in the paper, where in the desired angle $\psi_d(y_e)$ satisfies the requirement of the following formula (5)

$$\begin{cases} \psi_d(0) = 0 \\ y_e \psi_d(e) \geq 0, \forall y_e \in R \end{cases} \tag{5}$$

The expression of the desired angle $\psi_d(y_e)$ selected is the formula (6).

$$\psi_d = \arctan(k_e y_e), \qquad \psi_d \in (-\frac{\pi}{2}, \frac{\pi}{2}) \tag{6}$$

In equation (6), $k_e > 0$ is the normal tracking control parameter. It can be found from the above equation error dynamics equation that as long as the attitude tracking ψ_e angle converges to ψ_d and x_e converges to zero, then y_e naturally converges to zero, then it can be known that the approach angle ψ_d converges to zero.

Therefore, the control rate of the following formula (7) is selected in this paper. The speed of the reference point P is:

$$\begin{cases} U_p = U_d + k_\tau \tau_e \\ r_d = w_p - \dot{\beta} - \dot{\psi}_d + k_1(\psi_e - \psi_d) \\ \dot{s} = k_\tau \tau_e + v_t \cos\psi_e \end{cases} \tag{7}$$

Define the Lyapunov function as shown in the following equation (8)

$$V_1 = \frac{1}{2}(\psi_e - \psi_d)^2 \tag{8}$$

Its derivative is equation (9):

$$\begin{aligned}\dot{V}_1 &= (\psi_e - \psi_d)(\dot{\psi}_e - \dot{\psi}_d) = (\psi_e - \psi_d)(w_p - r - \dot{\beta} - \dot{\psi}_d) \\ &= -k_1(\psi_e - \psi_d)^2 \leq 0\end{aligned} \tag{9}$$

In the formula (9),k_1 is a number greater than zero. So V_1 is a monotonic non-increasing function, so function V_1 is bounded. Since V_1 has consistent continuity, it can be obtained according to Barbala's lemma:

$$\lim_{t\to\infty} \dot{V}_1 = 0 \Rightarrow \lim_{t\to\infty} \psi_e = \lim_{t\to\infty} \psi_d \tag{10}$$

Define the second Lyapunov function as shown in the following equation (11):

$$V_2 = \frac{1}{2}({\tau_e}^2 + {y_e}^2) \tag{11}$$

Derived for it, its derivative function is equation (12)

$$\begin{aligned}\dot{V}_2 &= \tau_e\dot{\tau}_e + y_e\dot{y}_e = \tau_e(\dot{s} - v_t\cos(\psi_e)) - y_e v_t \sin(\psi_e) \\ &= -k_2{\tau_e}^2 - y_e v_t \sin(\psi_e)\end{aligned} \tag{12}$$

Because $y_e\psi_d(y_e) \geq 0$,$V_t \geq 0$ and $y_e v_t \sin(\psi_e) \geq 0$, then V_2 is a monotonic non-increasing function. According to Barbala's lemma, it can be seen that x_e, y_e, ψ_e will converge to zero under the control rate of equation (7), so the robot fish will converge to the desired path.

3.2 Enhanced fireworks algorithm

The enhanced fireworks algorithm (EFWA) consists of an explosion operator, a mutation operator, a mapping rule, and a selection strategy. The flow chart of the enhanced fireworks algorithm is shown in Fig.3.

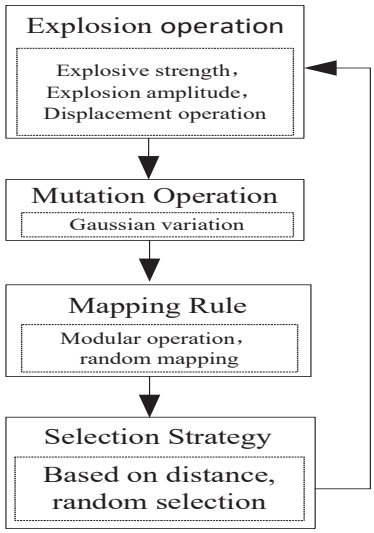


Figure 3: The flow chart of enhanced fireworks algorithm

(1) Explosion operator

$$S_i = SN \cdot \frac{y_{\max} - f_i + \varepsilon}{\sum_{i=1}^{N}(y_{\max} - f_i) + \varepsilon} \tag{13}$$

N is the total number of initial fireworks per generation in equation (13). S_i is the spark generated by each fireworks explosion. SN is the total number of explosion sparks.f_i is the individual fitness value, and y_max is the current population fitness value.S_i is a small number. The fireworks S_i have the following controls:

$$S_i = \begin{cases} round(a \cdot SN), S_i < a \cdot SN \\ round(b \cdot SN), S_i > b \cdot SN \\ round(S_i), \text{others} \end{cases} \tag{14}$$

a, b is a constant, and round is a rounding function. The following formula (15) is the magnitude of the explosion:

$$A_i = A \cdot \frac{f_i - y_{\min} + \varepsilon}{\sum_{i=1}^{N} (f_i - y_{\min}) + \varepsilon} \tag{15}$$

A is the maximum explosion amplitude of fireworks, and y_min is the optimal fitness value of the current population. $A_{\min,k}$ is the detection threshold with the lowest explosion radius in the kth dimension, that is:

$$A_{ik} = \begin{cases} A_{\min,k} \ A_{ik} < A_{\min,k} \\ A_{ik} \ others \end{cases} \tag{16}$$

where, t is the number of evaluations for the current iteration, $evals_max$ is the maximum number of evaluations, $A_{init}$and A_{final} are the initial and final blast radius detection values.

$$A_{\min,k}(t) = A_{init} - \frac{A_{init} - A_{final}}{evals_{\max}} \sqrt{(2evals_{\max} - t)t} \tag{17}$$

(2) Displacement operation and mapping rules The S_i explosion sparks are position offset according to equation (18) and then an explosion spark is generated.

$$\mathrm{X}_{ik} = X_{ik} + A_i \cdot U(-1, 1) \tag{18}$$

$U(-1, 1)$ is a uniform distribution between intervals $[-1, 1]$. When the spark x_i exceeds the boundary on the kth dimension, it will be out of bounds detection by the mapping rule of equation (19), and mapped to a new position.

$$X_{ik} = X_{LB,k} + U(0, 1) \times (X_{UB,k} - X_{LB,k}) \tag{19}$$

In equation (20),$X_{LB,k}, X_{UB,k}$ are the solution space at the lower and upper boundaries of the kth dimension, and $U(0, 1)$ is a uniformly distributed random number over the $[0, 1]$ interval.

(3) Gaussian variation The number N fireworks are randomly selected in kth dimensions, and each firework is subjected to Gaussian mutation operation in each dimension according to equation (20).

$$X_{ik} = X_{i,k} + (X_{B,k} - x_{ik})N(0, 1) \tag{20}$$

where, $N(0, 1)$ is a Gaussian distribution.

(4) Random selection strategy In order to ensure that the good individuals can be passed to the next generation in the iterative process, the optimal individual is first retained. And the remaining $N - 1$ fireworks are selected according to the random selection.

3.3 Design of EFWA-ADRC second-order auto-disturbance path tracking controller

It can be seen that each component of ADRC realizes functions independently by analyzing the principle and structure of ADRC. So the parameters are independently set according to the principle of separation. The design of the auto disturbance rejection controller is simplified [23]. Among them, for the selection of the nonlinear parameter,$\beta_1, \beta_2, \beta_{01}, \beta_{02}, \beta_{03}$, an enhanced fireworks algorithm is adopted. In order to obtain satisfactory control performance, the requirements of system speed, stability, accuracy and control quantity are considered comprehensively in this paper. The ITME criterion is selected as the fitness function to evaluate the system performance. The enhanced fireworks algorithm is used to optimize the parameters of ADRC. It avoids the complicated trial and error times of ADRC in the actual debugging process. Forward speed control is taken as an example, the flowchart of the ADRC controller based on the enhanced fireworks algorithm constructed in this paper is shown in Fig.5.

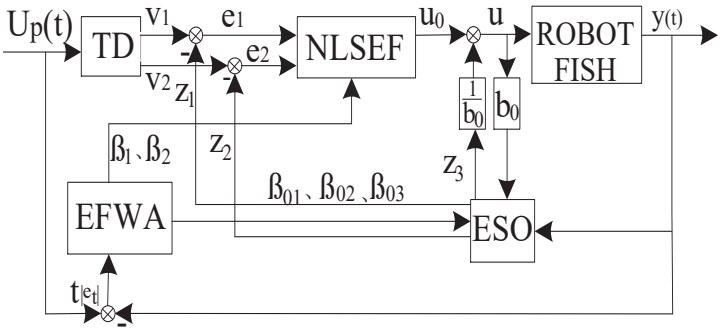


Figure 4: Structure diagram of forward speed controller based on ARDC

In Fig.4, $U_{p(t)}$ is the input signal (the desired speed of the robot fish). $v_1(t)$ is the transition process for extracting the $U_{p(t)}$ signal from TD. And v_2 is the differential signal obtained through. $y(t)$ is the output signal of the controlled object. z_1, z_2 are the state variables of the controlled object of ESO real-time estimation. z_3 is the disturbance of the controlled object estimated by ESO in real time. e_1, e_2 is the error signal corresponding to v_1, v_2 respectively. $u_0(t)$ is the initial control quantity of the controlled object obtained by NLSEF. $u(t)$ is the final control quantity after real-time compensation disturbance. And b is the compensation factor. The $u(t)$ and $y(t)$ subtractive, then absolute errors are obtained, multiplied by the time t to construct the fitness function. After EFWA optimization calculation, the nonlinear parameters β_1,β_2,β_{01},β_{02},β_{03} are adjusted online to optimize the performance of the system.

The design of the steering speed controller is similar to the design method of the forward speed controller. The paper purpose is to find the optimal control rate F_u and tau through the ADRC controller. So the tangential and normal errors are tended to zero. Thereby the robot fish converges to the desired path. Fig.5 is a flow chart based on the WA-ADRC robotic fish path tracking controller. The desired error information τ_e, η_e, ψ_e can be solved by the desired path information, the real-time position of the robot fish. And the desired control rate U_p,r_d can be generated from the error informationτ_e, η_e, ψ_e . The control rates U_p,r_d are input to the forward speed and steering speed controllers respectively, and the control information F_u and tau are generated by the controller. Considering that the robotic fish is an under-actuated model, the lateral input of the robot fish is $F_v = 0$. And the disturbance is input into the robot fish. After the path generator, the robot fish real-time position is finally obtained. It is obtained the second-order EFWA-ADRC path tracking controller, so that the problem of the position error

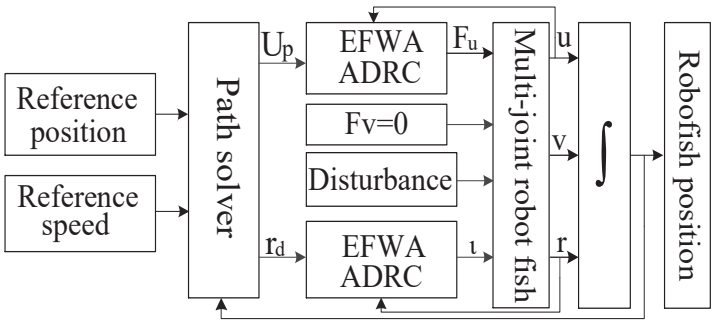


Figure 5: Structure diagram of EFWA-ADRC path tracking controller

stabilization of the robot fish is converted into the stabilization problem of the speed and the heading.

4 Simulation analysis and experimental verification

The quality of the robot fish in the paper is$M = 8kg$, full length is $I_B = 0.640m$, the length of trunk is$0.440m$, the length of the stalk is$0.080m$, the length of the caudal fin is$0.120m$, $I_x = 0.1095kgm^2, I_y = 0.4637kgm^2, I_z = 0.4472kgm^2$, the viscosity coefficient of water movement is$v = 1 \times 10^6 m^2/s$. In order to verify the performance of the self-disturbance path tracking controller proposed in the paper, according to the model of the above-mentioned robot fish. The robot fish path tracking simulation system is built in the SIMULINK environment. The linear and circular tracking tasks were used to design simulation experiments and tracking experiments to verify the performance of the path tracking designed and compare it with the conventional ADRC and PID control methods.

4.1 Simulation analysis of linear tracking

(1) Simulation analysis of linear tracking Give the linear parameter equation that the robot fish expects to track as equation (23):

$$\begin{cases} \xi_{\mathrm{p}}(s) = \sin\theta \\ \eta_p(s) = \cos\theta \end{cases} \tag{21}$$

Where $\theta = \frac{\pi}{4}$is the angle between the straight path and the x-axis. The initial value of the reference path is $(0, 0)$. The expected speed is$U_d = 1m/s$. The initial position of the robot fish is$(0, 0)$,$\psi = 0$. The initial speed is 0.

Fig.6 shows that the tracking system designed has stability in the tracking task of the straight path. And all three controllers complete the tracking task. It can be seen from Fig.7 that the enhanced fireworks algorithm basically finds the global optimal value after 60 iterations. The fitness function is basically maintained at about 10. In terms of speed control of the robot fish, it can be seen the forward speed and steering speed of the EFWA-ADRC controlled robot can reach the set value in $3s$, the ADRC is $3.2s$ in Fig.8 (a) and (b), and in contrast, PID control converges slowly and stays near the expected value at $4s$. It can be seen the time taken by the three control methods of EFWA-ADRC, ADRC, and PID is $3s$, and $4s$ respectively in order to keep the position error of the robot fish positive and negative errors of $0.1m$ from Fig.9 (a), (b), and (c). From Fig. 6 to Fig. 9, it can be seen that the controller based on EFWA-ADRC has

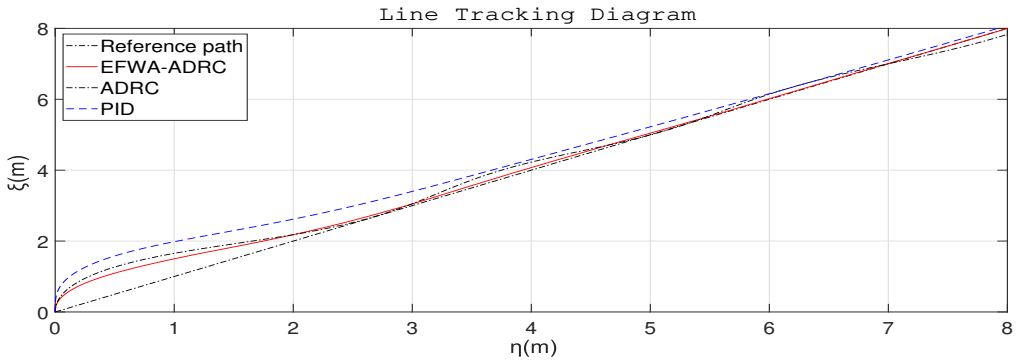


Figure 6: Curve path trace diagram of horizontal plane

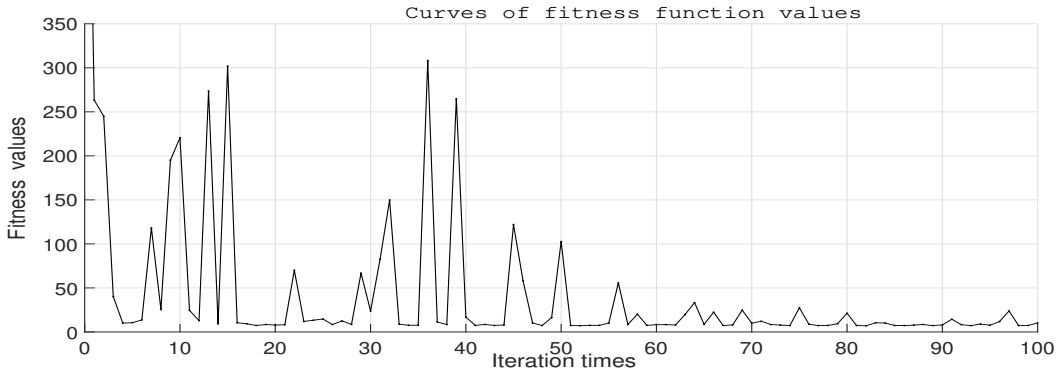
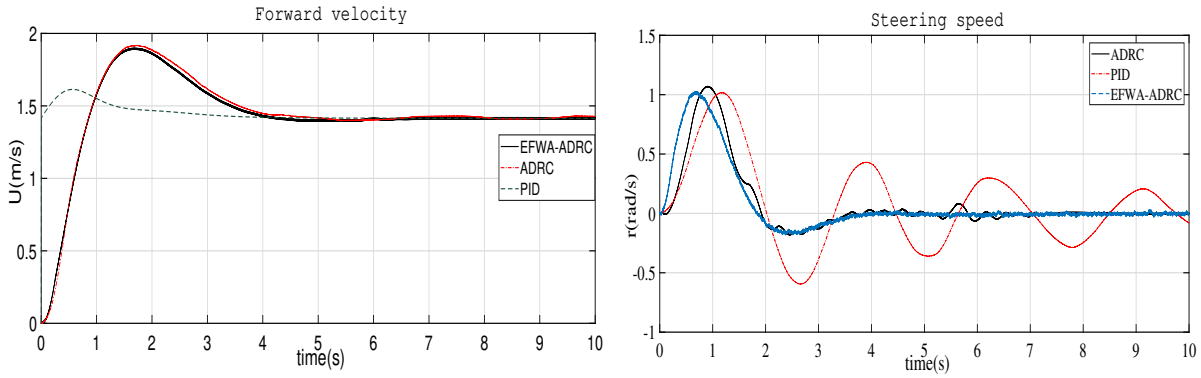


Figure 7: Fitness function curves

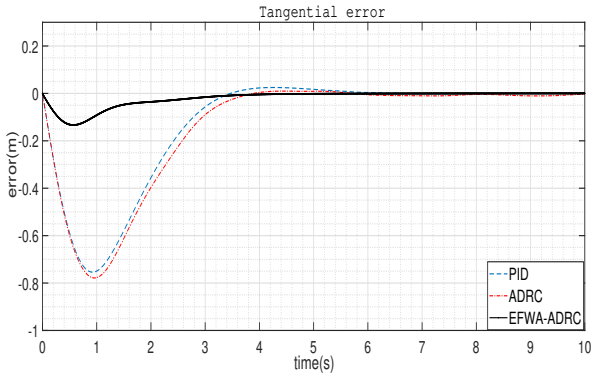


(a) Forword speed. (b)Steering speed.

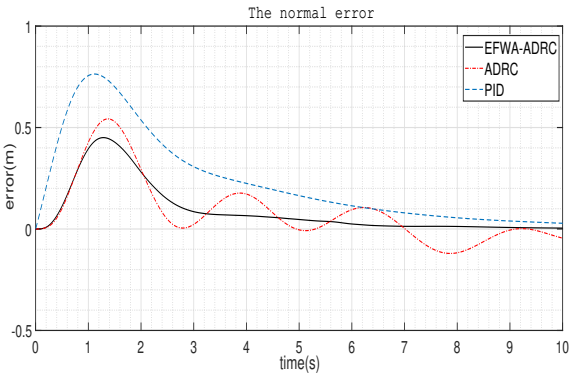
Figure 8: Robot fish speed curves

short transition time, small overshoot and good anti-interference performance during tracking. Compared with the ADRC control method, the speed response time and position error of the robot fish have been improved, and the control accuracy has been improved by 10%. Therefore, the path tracking controller based on EFWA-ADRC has better dynamic performance and static performance in the linear tracking process is proved.

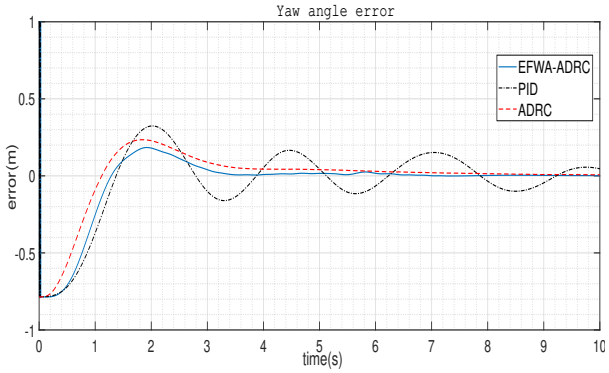
(2) Curve path tracking simulation analysis It is $\Gamma(s)$ the parameter equation for the robotic fish to track the curve path.



(a) Tangential error.



(b)Normal error.



(c) ψ angular error.

Figure 9: Straight line tracking error curves

$$\begin{cases} \xi_p(s) = \cos(s) \\ \eta_p(s) = \sin(s) \end{cases} \tag{22}$$

The initial value of the reference curve path is $(0, 0)$. A circle with a radius of $1m$, clockwise rotation, the expected speed is $U_d = 1m/s$. The initial position of the robot fish is $(0, 0)$. $\psi = 0$, the initial speed is0. The simulation results are shown in Fig.10 to Fig.14.

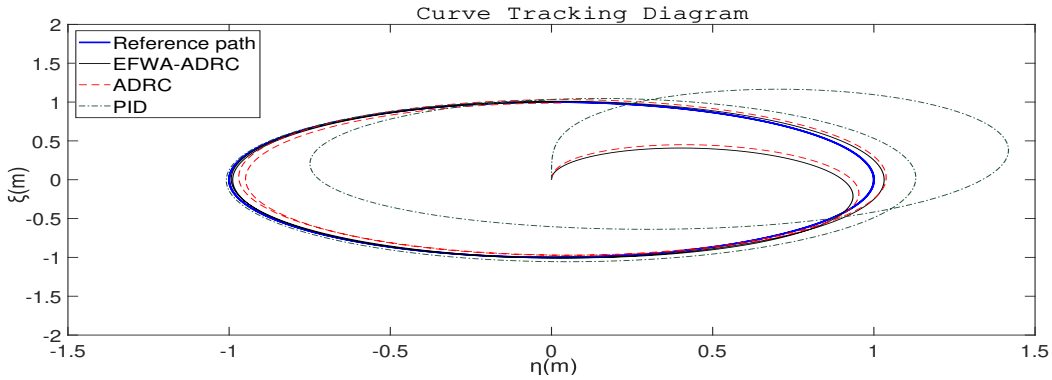


Figure 10: Curve path trace diagram of horizontal plane

It can be seen from Fig. 10 that the ADRC controller has strong robust performance. The tracking system designed also has good stability to the curve tracking task. It can make the

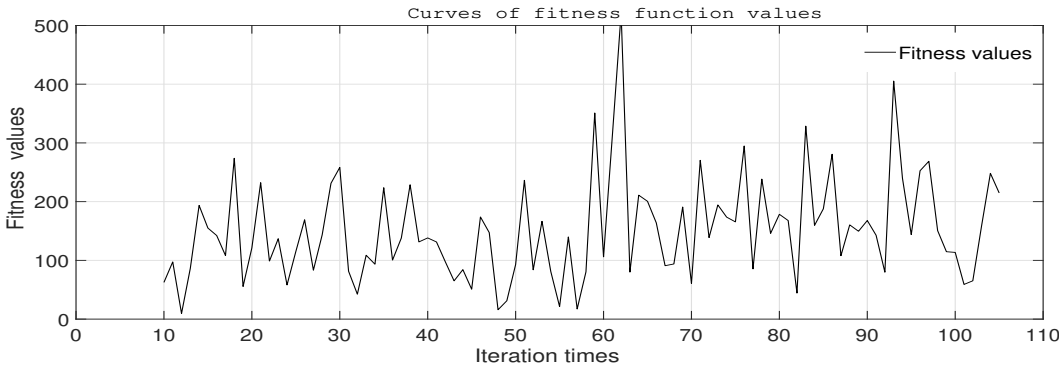
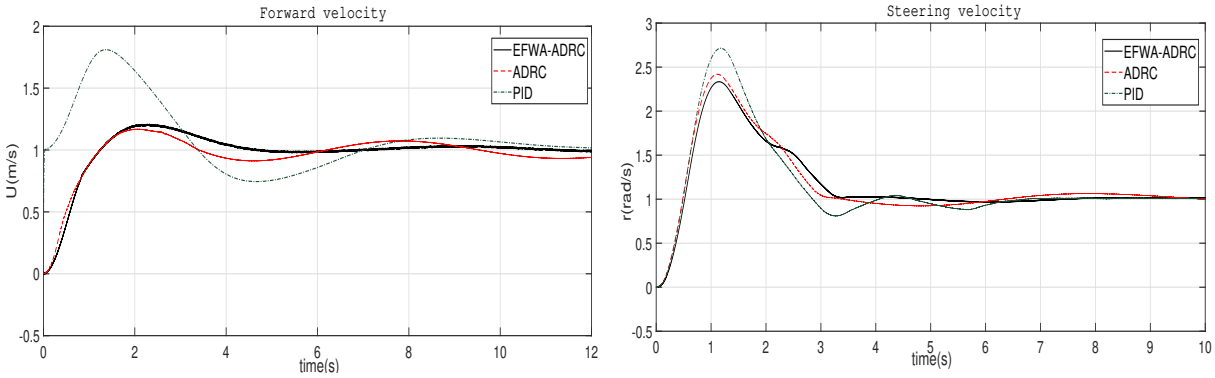


Figure 11: Fitness function curves



(a) Forword speed. (b)Steering speed.

Figure 12: Robot fish speed curves

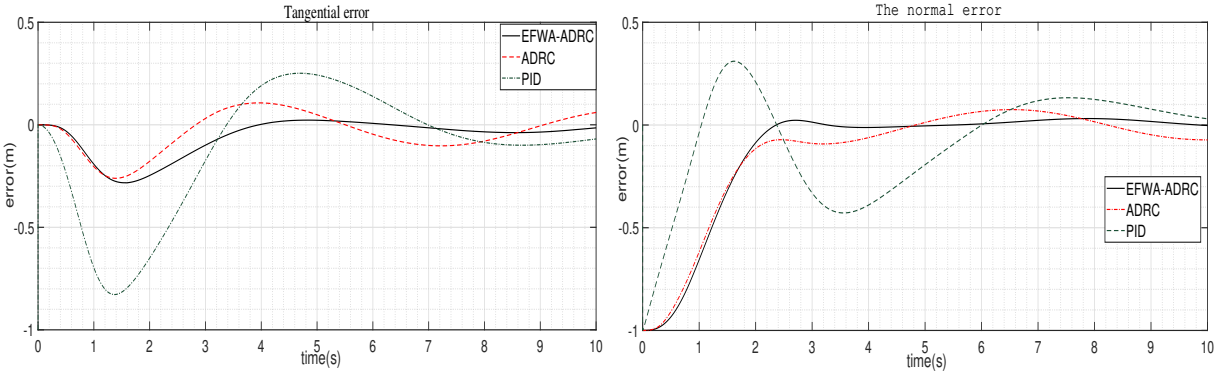
robot fish converge to the desired path. In Fig.11 the optimal value of the fitness function of the system reaches 5.21. It can be seen from Fig. 12 that the forward speed and steering speed of the EFWA-ADRC controlled robot can reach the set value in $3s$, the ADRC is $3.5s$, and the PID method converges slowly, about $4s$. It can be seen from Fig. 13 that the time taken by the three control methods of EFWA-ADRC, ADRC, and PID is $2.8s$, $3s$, and $4s$ respectively in order to keep the position error of the robot fish positive and negative errors of$0.1m$. It can be concluded from Fig. 10 to Fig. 13 that the fastness and robustness based on the EFWA-ADRC controller are satisfactory.

Based on the above simulation experiments, it is proposed that the self-disturbance path tracking controller has good control accuracy. Compared with the conventional ADRC controller, the EFWA-ADRC control method has good control effect in terms of accuracy and response time, and it has high control quality and control precision.

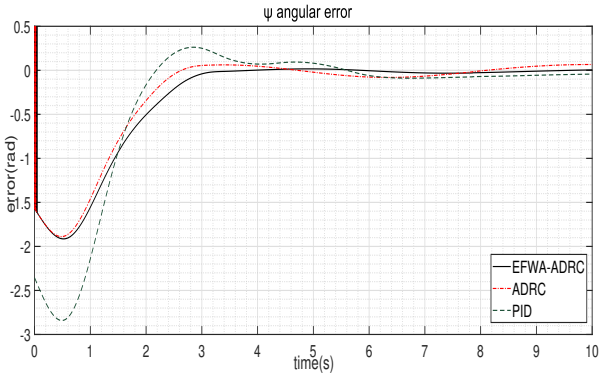
4.2 Experimental verification

The experimental experiments were carried out on the underwater experimental platform to verify the feasibility of the proposed method. The four-joint robot fish used in the experiment. The length is $0.64m$ and the weight is$8kg$. The size of the fish pond is and the mass of the robot fish is $8Kg$.

(1) Straight path tracking experiment verification The initial position of the robot fish is $(10, 5)$, $\psi = 0$the initial speed is 0m/s, the reference path is a straight line from $(0, 0)$ to $(300, 200)$,



(a) Tangential error. (b)Normal error.



(c) ψ angular error.

Figure 13: Curve tracking error curves

the expected speed of the robotic fish in the experiment is $1m/s$. it showed in the fig.14- fig.18 the linear path tracking experiment process. In order to reduce the complexity of the debugging process, the ADRC path tracking method with optimized parameters and the tracking method based on the traditional PID are compared. is the experimental picture is showed in the fig.14 based on a conventional PID control method. Fig.15 is an experimental diagram based on the EFWA-ADRC path tracking method.

(2) Curves path tracking experiment verification

In the curve path tracking experiment, the robot fish initial position is $(150, 120)$,$\psi = 0$. The initial speed is $0m/s$. the expected speed is $1m/s$. The reference path is a circle with a center of $(150, 120)$ and a radius of $50cm$. The starting point of the reference path is $(200, 170)$. The direction angle is$\psi = \pi/2$. Fig.16 is an experimental diagram based on a conventional PID control method. Fig.17 is an experimental diagram based on the EFWA-ADRC path tracking method. Fig.18 (a) shows the positional trajectory of the robot fish during the linear path tracking process. The yellow line represents the robotic trajectory of the traditional PID control method, and the red one represents the EFWA-ADRC control robotic trajectory. Fig.18(b) is the trajectory during the curve path tracking process, and the desired path represented by black in the two figures.

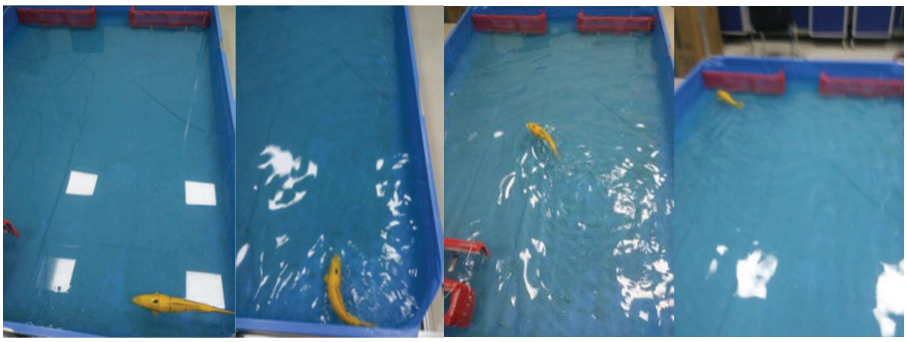


Figure 14: Experimental pictures based on traditional PID control method

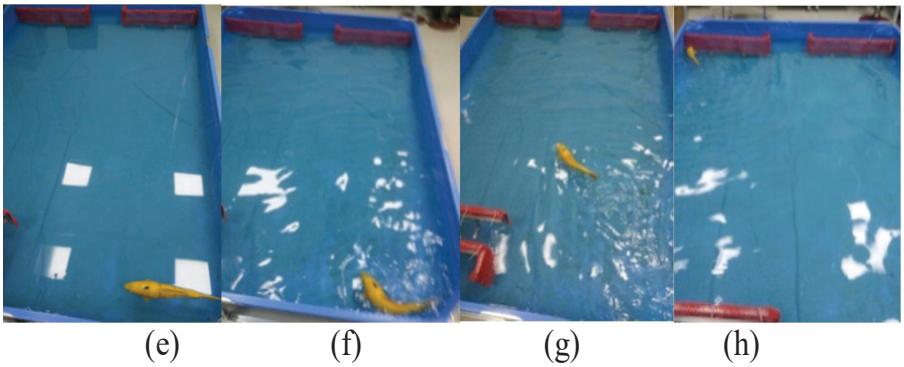


Figure 15: Experimental pictures based on the EFWA-ADRC path tracking method

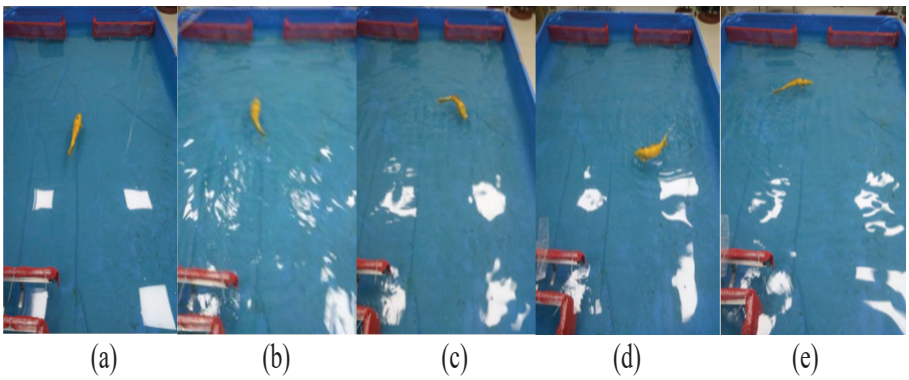


Figure 16: Experimental pictures based on traditional PID control method

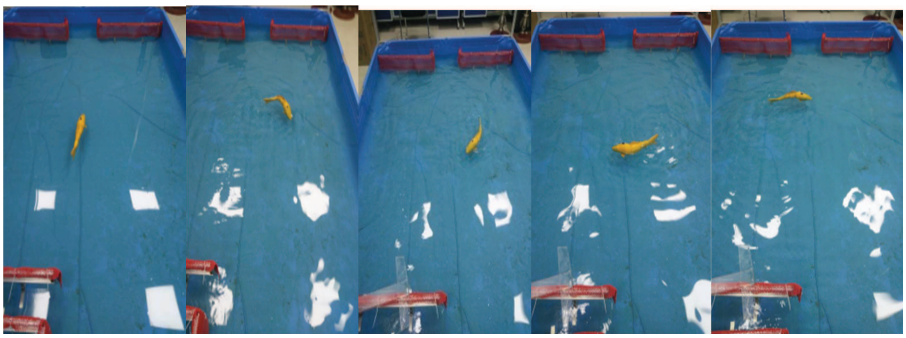


Figure 17: Experimental pictures based on the EFWA-ADRC path tracking method

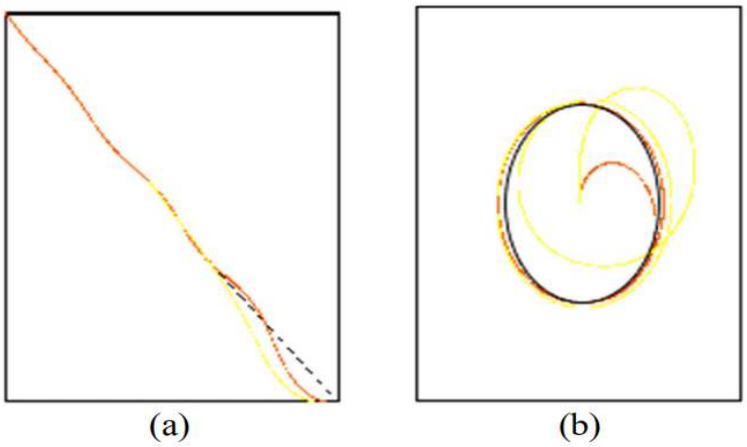


Figure 18: The experiment trajectories of the robotic fish

The yellow trajectory in Fig.18 shows that the robot fish trajectory is highly volatile due to the interference of water waves. There is no transition process in the PID control method. It results in a large overshoot of the speed control, and the position error fluctuates greatly. In the ADRC controller the transition process is arranged. So the robot fish speed changes are relatively flat. The unknown disturbance can be estimated and compensated. The red trajectory also proves that the ADRC controller has good anti-interference performance. Therefore, from the combination of simulation and experimental results, the feasibility of the proposed method is proved. The control system has better control precision.

5 Conclusion

The robot fish path tracking problem is mainly studied in the paper. The robotic fish path tracking controller is proposed based on the EFWA-ADRC. The virtual mobile robot is introduced combining the kinematics and dynamics equations of multi-joint robotic fish. The tracking error model of the curve path is established in the Serret-Frenet coordinate system. The guidance function of forward and steering control is designed. The tracking errors were reduced by adjusting path parameters and corner velocity. The second-order path controller based on EFWA-ADRC is established. The control information f and ? are generated. Ultimately the robotic fish can track a given route at a certain speed. Moreover the ADRC parameters are optimized by the fireworks algorithm. So the control precision is further improved. Through simulation analysis and experiment, the feasibility of the proposed method is verified. It has good adaptability to systems with interference and uncertainty. It is further proved that the EFWA-ADRC path tracking controller designed has good control quality in the paper.

Funding

This project is grateful for the following projects: National Key R&D Program (2016YFE0111900), Shaanxi International Science and Technology Cooperation Program (2018KW-022), Shaanxi Natural Science Foundation (2018JQ5009 and 18JK0388), and the independent intelligent control research and innovation team support this topic.

Conflict of interest

The authors declare no conflict of interest.

Bibliography

[1] Bak, J.; Nguyen, H.N.; Park, S.; Lee, D.; Seo, T.W.; Jin, S.; Kim, J. (2017). Positioning control of an underwater robot with tilting thrusters via decomposition of thrust vector, *International Journal of Control Automation & Systems*, 15(5), 1-9, 2017.

[2] Barrett, D.S.; Triantafyllou, M.S., Yue, D.K.P.; Grosenbaugh, M.; Wolfgang, M.J. (1999). Drag reduction in fish-like locomotion, *Journal of Fluid Mechanics*, 192(392), 183-212, 1999.

[3] Bian, X.Q.; Cheng, X.Q.; Jia, H.M.; Yan, Z.P.; Zhang, L.J. (2011). A bottom-following controller for underactuated AUV based on iterative sliding and increment feedback, *Control & Decision*, 26(2): 289-292, 296, 2011.

[4] Caharija, W.; Pettersen, K.Y.; Bibuli, M.; Calado, P.; Zereik, E.; Braga, J.; Gravdahl, J.T.; Sorensen, A.; Milovanovic, M.; Bruzzone, G. (2016). Integral Line-of-Sight Guidance and Control of Underactuated Marine Vehicles: Theory, Simulations, and Experiments, *IEEE Transactions on Control Systems Technology*, 24(5), 1623-1642, 2016.

[5] Chao, D.; Yin, Z.; Zhang, Y.; Liu, J.; Sun, X.D.; Zhong, Y.R. (2019). Research on Active Disturbance Rejection Control of Induction Motors Based on Adaptive Particle Swarm Optimization Algorithm with Dynamic Inertia Weight, *IEEE Transactions on Power Electronics*, 34(3), 2841-1855, 2019.

[6] Fan, S.; Li, B.; Xu, W.; Xu, Y. (2018). Impact of Current Disturbances on AUV Docking: Model-Based Motion Prediction and Countering Approaches. *IEEE Journal of Oceanic Engineering*, 43(4), 888-904, 2018.

[7] Fossen, T.I.; Lekkas, A.M.(2017). Direct and indirect adaptive integral line-of-sight path-following controllers for marine craft exposed to ocean currents. *International Journal of Adaptive Control & Signal Processing,*, 31(4), 445-463, 2017.

[8] Gao, C.; Yuan, J.; Zhao, Y. (2018). ADRC for spacecraft attitude and position synchronization in libration point orbits. *Acta Astronautica*, 145, 238-249, 2018.

[9] Haga, M.; Uchida, M. (2017). Propulsion modeling of caudal fin driving system on balloon fish robot, *Artificial Life & Robotics*, 22(1), 10-16, 2017.

[10] Han, J.Q. (1998). Auto-disturbance Controller and Its Application, *Control and Decision*, 1998(1), 19-23, 1998.

[11] Jung, S. (2018). Improvement of Tracking Control of a Sliding Mode Controller for Robot Manipulators by a Neural Network, *International Journal of Control Automation & Systems*, *IEEE Int. Conf. on Robotics and Automation*, 16(2), 937-943, 2018.

[12] Kelasidi, E.; Liljeback, P.; Pettersen, K.Y. et al(2017). Integral Line-of-Sight Guidance for Path Following Control of Underwater Snake Robots: Theory and Experiments, *IEEE Transactions on Robotics*, 33(3), 610-628, 2017.

[13] Miller, T.F. (2018). A bio-inspired climb and glide energy utilization strategy for undersea vehicle transit, *Ocean Engineering*, 149, 78-94, 2018.

[14] Peng, Y.Y.; Yi, G. (2018). Research on Exploring the Patients? Hiding Disease Based on Symptom Weighted Clustering Technique. *International Journal Bioautomation*, 22(3), 263-274, 2018.

[15] Song, X.R. (2018). A New Hybrid Method in Global Dynamic Path Planning of Mobile Robot, *International Journal of Computers Communications & Control*, 13(6), 1032-1046, 2018.

[16] Song, X.; Chen, H.; Xue, Y. (2015). Stabilization precision control methods of photoelectric aim-stabilized system, *Optics Communications*, 351, 115-120, 2015.

[17] Song, X.R.; Gao, S.; Chen, C.B. (2018). A Novel Vehicle Feature Extraction Algorithm. *Traitement Du Signal*, 35(3-4), 223-242, 2018.

[18] Tan, Y.; Zhu, Y. (2010). Fireworks Algorithm for Optimization, Advances in Swarm Intelligence. *Springer Berlin Heidelberg*, 355-364, 2010.

[19] Wu, L.; Bao, H.; Du, J.L.; Wang, C.S. (2014). A learning algorithm for parameters of automatic disturbances rejection controller. *Acta Automatica Sinica*, 40(3), 556-560, 2014.

[20] Xiao, Q.; Wang, Y.; Wang, H.(2015). Motion retrieval using weighted graph matching. *Soft Computing*, 19(1), 133-144,2015.

[21] Xiao, Q.K.; Liu, S.Q. (2017). Motion retrieval based on Dynamic Bayesian Network and Canonical Time Warping, *Soft Computing*, 21(1), 267-280, 2017.

[22] Xue, P.; Jiang, C.H.; Wei, W.; Lin, J. (2018). Optimization of the intelligent workshop control based on the improved group leadership optimization algorithm. *International Journal of Simulation Modelling,*, 17(4), 690-702, 2018.

[23] Yang, J.; Ding, Z.; Chen, W.H.; Li, S.H. (2016). Output-based disturbance rejection control for non-linear uncertain systems with unknown frequency disturbances using an observer backstepping approach, *Iet Control Theory & Applications*, 10(9), 1052-1060, 2016.

[24] Yen, W.K.; Sierra, D.M.; Guo, J. (2018). Controlling a Robotic Fish to Swim Along a Wall Using Hydrodynamic Pressure Feedback, *IEEE Journal of Oceanic Engineering*, 43(2), 369-380, 2018.

[25] Zhou, X.; Gao, H.; Zhao, B.; Zhao, L. (2018). A GA-based parameters tuning method for an ADRC controller of ISP for aerial remote sensing applications, *ISA Transactions*, 81, 318-328, 2018.

[26] Zhang, G.C. ; Huang, H.; Qin, H.D. et al(2017). A novel adaptive second order sliding mode path following control for a portable AUV, *Ocean Engineering*, 151, 82-92,2018.

INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL
ISSN 1841-9836, e-ISSN 1841-9844, 14(3), 419-436, June 2019.

Reliability Assessment Model for Industrial Control System Based on Belief Rule Base

Y.H. Wang, P.L. Qiao, Z.Y. Luo, G.L. Sun, G.Z. Wang

Yuhe Wang, Peili Qiao*, Zhiyong Luo, Guanglu Sun, Guangze Wang
School of Computer Science and Technology
Harbin University of Science and Technology
Harbin 150080, China
*Corresponding author: qiaopl@hrbust.edu.cn

Guangze Wang
Library
Harbin University of Science and Technology
Harbin 150080, China

Abstract: This paper establishes a novel reliability assessment method for industrial control system (ICS). Firstly, the qualitative and quantitative information were integrated by evidential reasoning(ER) rule. Then, an ICS reliability assessment model was constructed based on belief rule base (BRB). In this way, both expert experience and historical data were fully utilized in the assessment. The model consists of two parts, a fault assessment model and a security assessment model. In addition, the initial parameters were optimized by covariance matrix adaptation evolution strategy (CMA-ES) algorithm, making the proposed model in line with the actual situation. Finally, the proposed model was compared with two other popular prediction methods through case study. The results show that the proposed method is reliable, efficient and accurate, laying a solid basis for reliability assessment of complex ICSs.
Keywords: Belief rule base (BRB), industrial control system (ICS), evidential reasoning (ER), reliability assessment, covariance matrix adaptation evolution strategy (CMA-ES) algorithm.

1 Introduction

The industrial control system (ICS) integrates computer technology, communication and the control theory [9]. It is widely used in infrastructure industries like power transmission, water supply, oil and gas transportation, etc. In the ICS, each controller regulates multiple components. The fault or intrusion of one component will threaten the security of the entire system. Once the ICS fails, the user will not only lose monitoring and control, but also suffer from facility damage, economic losses and even casualties. Therefore, it is of great significance to accurately assess the reliability of the ICS, especially in complex applications [12].

The existing ICS reliability assessment methods are mainly based on knowledge, model or data [18]. The knowledge-based methods include fault mode, effects, and criticality analysis (FMECA) [10], fault tree [8], decision tree [1], and risk analysis [15]. These approaches mainly rely on qualitative or quantitative knowledge. The model-based methods, namely, open-switch fault diagnosis [20], signal-based coding [11] and processing fault diagnosis [12], can diagnose the faults of different systems according to the actual industrial processes. However, it is difficult to build effective models for large-scale ICSs. The data-based methods emerge due to the proliferation of industrial data collection techniques, such as distributed control system (DCS) and supervisory control and data acquisition (SCADA). Relying on historic process data, many data-based methods are suitable for ICS fault detection, e.g. Shewhart individuals control chart [5], Hotelling's T-squared (T^2) control chart [2], quality control chart [17], principal component analysis (PCA) [16], and knowledge discovery in databases (KDD) [19].

Each type of the above methods has its defects. For knowledge-based methods, the experts are often unable to obtain the accurate qualitative knowledge, owing to the complexity of the ICS and the numerous factors [13]. The model-based methods depend heavily on specific samples and cannot be extended easily to general cases. Neither can they use qualitative or quantitative data. The data-driven methods work well in reliability assessment of accurate data samples, but perform poorly in differentiating normal data from abnormal data [14]. What is worse, the ICS data cannot be acquired under certain conditions. To sum up, the existing methods cannot effectively utilize all the various uncertain information in the ICS, including expert knowledge and historical data. It is imperative to develop an approach to assess ICS reliability against these semi-quantitative data.

To solve the above problems, this paper integrates qualitative and quantitative information by evidential reasoning (ER) rule, and then establishes an ICS reliability assessment model based on belief rule base (BRB), which is a powerful nonlinear strategy for uncertainty problems [6] [22]. In this way, both expert experience and historical data were fully utilized in the assessment. Finally, the initial parameters were optimized by covariance matrix adaptation evolution strategy (CMA-ES) algorithm, making the proposed model in line with the actual situation [7].

The remainder of this paper is organized as follows: Section 2 analyzes the reliability of the ICS; Section 3 sets up a BRB-based model and applies it to assess the ICS reliability; Section 4 compares the proposed model with two other methods through a case study on a tobacco factory; Section 5 wraps up this paper with several meaningful conclusions.

2 Problem overview

2.1 ICS structure

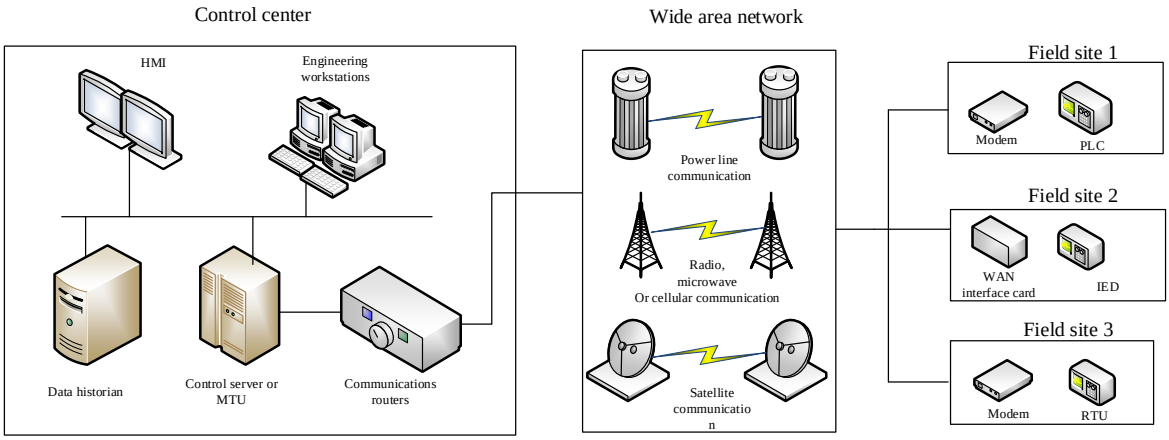


Figure 1: Typical SCADA structure

Consider SCADA as a typical ICS, the structure of SCADA is illustrated in Figure 1 [21]. In the control center, there are a human-machine interface (HMI), an engineering workstation, a data historian, a master terminal unit (MTU) and a communication router. All these components are connected via a local area network (LAN). Among them, the HMI mainly displays the data collected from the site, while the MTU stores and processes input and output data. In the wide area network, the information is transmitted between the control center and the field sites via power line, radio, microwave, cellular or satellite. On the field sites, there are multiple distributed remote terminal units (RTUs) or programmable logic controllers (PLCs), which control local

processes. The ICS can collect various data, analyze the trend and issue alarms when parameters exceed the allowable range.

According to the struture of SCADA, the main components of ICS can be divided into three types as shown in Table 1 [3].

Table 1: The type of main components in ICS

First level	Second level	Third level
ICS components	Hardware components	PLC
		RTU
		HMI
		MTU
		Other hardware
	Software components	OS(Operation system)
		DB(Data base)
		Software of SCADA system
		Software of PLC
		Other software
	Communication	Profinet
		Communication routers
		Modem
		Other communication components

2.2 The framework for ICS reliability assessment

The accuracy of ICS reliability assessment hinges on the clear identification of various influencing factors. In fact, these factors can be divided into internal factors and external factors. As shown in Table 1, the internal factors are the faults of the system components, including software fault, hardware fault and communication fault, while the external factors include network attacks and human errors, both of which directly bears on the system security. Therefore, the ICS reliability can be assessed from both the fault and security of the system.

On this basis, the framework of ICS reliability assessment was constructed. As shown in Figure 3, the framework consists of three layers: the target layer, the criteria layer and the plan layer. The target layer is the reliability assessment; the criteria layer encompasses fault assessment and security assessment; the plan layer contains various antecedent attributes that affect the criteria layer.

2.3 Process of ICS reliability assessment

Firstly, the fault and security attributes of the ICS were integrated by the ER rule for fault and security analyses. Based on the analysis results, the BRB was adopted to establish a model and CMA-ES for training. Finally, the ICS reliability assessment model was obtained. The specific steps are presented in Figure 4 below.

2.4 Mathematical description

Let FA, SC and FR be the fault assessment result, security assessment result and final result of ICS reliability assessment, respectively. The set of antecedent attributes C^1 and that C^2 for fault assessment and security assessment can be respectively expressed as:

$$C^1 = \left\{c_1^1, c_2^1, \cdots c_n^1\right\} \tag{1}$$

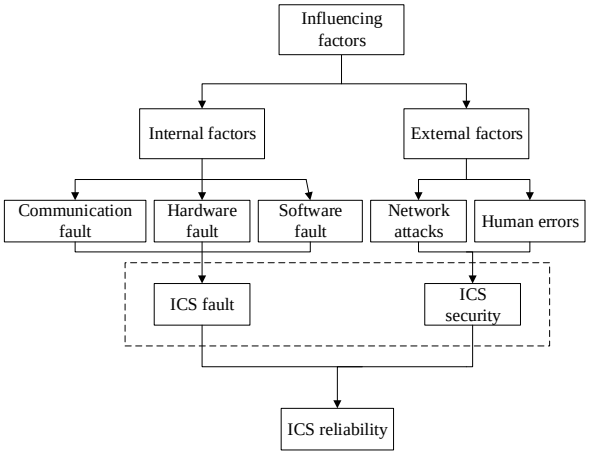


Figure 2: Relationship between influencing factors and ICS reliability

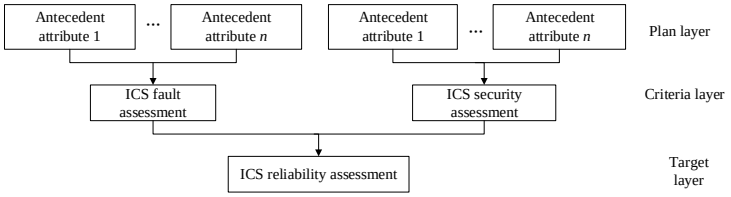


Figure 3: Framework of ICS reliability assessment

$$C^2 = \left\{c_1^2, c_2^2, \cdots c_m^2\right\} \tag{2}$$

where c_i^1 and c_i^2 are the i-th antecedent attribute for fault assessment and security assessment, respectively.

Then, the fault assessment result and security assessment result can be respectively obtained by:

$$FA = ER_1\left(C^1, t^1\right) \tag{3}$$

$$SC = ER_2\left(C^2, t^2\right) \tag{4}$$

where $ER_1(\bullet)$ is the relationship between the leading attribute and the fault assessment result; t^1 is the set of parameters for $ER_1(\bullet)$; $ER_2(\bullet)$ is the relationship between the leading attribute and the security assessment result; t^2 is the parameters of $ER_2(\bullet)$. On this basis, the BRB-based ICS reliability assessment result can be obtained by combining the fault and security assessment results:

$$FR = BRB(FA, SC, t) \tag{5}$$

where $BRB(\bullet)$ is the conversion from fault and security assessment results to reliability assessment result; t is a set of parameters.

3 Construction of BRB-based ICS reliability assessment model

As mentioned before, the first step of ICS reliability assessment is to clearly identify all the faults in the system components. After all, the fault of one component will threaten the security of the entire system. Moreover, the ICS reliability assessment must consider both quantitative data (e.g. system duration, fault frequency and PLC fault tolerance) and qualitative data (e.g.

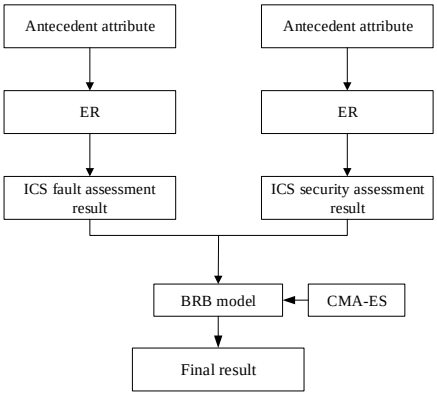


Figure 4: The process of ICS reliability assessment

terminal controllability, communication stability). The former can be collected by data monitoring system, while the latter require expert interpretation. If the two types of data are uncertain, it is impossible to assess the ICS reliability in an accurate manner.

Considering the excellent performance of the ER and the BRB for uncertainty problems, the author decided to set up a multi-layer ICS reliability assessment model, optimize the model with the historical data of the BRB, and integrate multiple attributes with the ER rule. The proposed model contains a fault assessment model and a security assessment model. Besides, the parameters was trained by the CMA-ES algorithm, such that the assessment indices are in line with the actual situation.

3.1 Fault assessment attributes

Reliability is defined as the ability or possibility to perform a specified function without failure within a specified time and under specified conditions. The reliability of the system can be evaluated by mean time to repair, failure rate and mean time to failure [22].

In our framework of ICS reliability assessment, the fault assessment covers three aspects. For each attribute, the weight t is the most important parameter in the ER process. The antecedent attributes of fault assessment are listed in Table 2.

3.2 Security assessment attributes

The ICS security is partially affected by network attacks. The common attacks include virus and trojan attack, DoS attack, detection attack, U2R attacks and R2L attack. Specifically, the virus and trojan attack destroys the ICS or steals system data with a malicious program; the DoS consumes an excessive amount of ICS resources, making services unavailable; the detection attack scans the ICS network and unlocks the hidden security dangers; the U2R attack acquires the rights of the ICS superuser through manipulation of the ICS vulnerabilities; the R2L attack remotely gains unauthorized access to the ICS. The different attacks pose varied levels of threats and brings diverse damages to the ICS. Therefore, the attack type, continuous attack time, attack frequency and attack severity were selected as the antecedent attributes of the security assessment.

The ICS security is also affected by human errors like mis-operation, unauthorized entry, etc. Hence, the occurrence frequency and severity of such errors were also taken as the antecedent attributes of the security assessment. Table 3 lists all the antecedent attributes of the security assessment.

Table 2: The antecedent attributes of fault assessment

First level	Second level	Third level
The Fault Reliability of ICS	Hardware Reliability of ICS	The Rate of Failure $(c_{11}^1)(t_{11}^1 = 0.15)$
		MTTF(Mean Time to Failure $(c_{12}^1)(t_{12}^1 = 0.2)$
		MTTR(Mean Time to Repair) $(c_{13}^1)(t_{13}^1 = 0.15)$
		MPMT(Mean Preventive Maintenance Time) $(c_{14}^1)(t_{14}^1 = 0.1)$
		MPBF(MeanTime Between Failure) $(c_{15}^1)(t_{15}^1 = 0.1)$
		The Failure Severity $(c_{16}^1)(t_{16}^1 = 0.15)$
		The Failure Tolerance $(c_{17}^1)(t_{17}^1 = 0.15)$
	Software Reliability of ICS$(c_2^1)(t_2 = 0.15)$	MTTF(Mean Time to Failure) $(c_{21}^1)(t_2^1 = 0.2)$
		MTTR(Mean Time to Repair) $(c_{22}^1)(t_{22}^1 = 0.2)$
		The Fluency of Software $(c_{23}^1)(t_{23}^1 = 0.3)$
		MPBF(Mean Time Between Failure) $(c_{24}^1)(t_{24}^1 = 0.15)$
		The Rate of Failure $(c_{25}^1)(t_{25}^1 = 0.15)$
	Communication Reliability of ICS	The Rate of Lost $(c_{31}^1)(t_{31}^1 = 0.2)$
		The Packet Loss Rate $(c_{32}^1)(t_{32}^1 = 0.3)$
		MTTF(Mean Time to Failure) $(c_{33}^1)(t_{33}^1 = 0.2)$
		The Failure Severity $(c_{34}^1)(t_{34}^1 = 0.2)$
		Delay rate $(c_{35}^1)(t_{35}^1 = 0.15)$

3.3 ER rule process

ER rule is developed as a multi-criteria decision analysis (MCDA) approach on the basis of belief decision and D-S theory [22]. Compared to the D-S theory of evidence, the calculation process of the ER rule is linear. The ER rule can be implemented in the following steps:

Assume that there are P basic attributes $\{c_1, \cdots c_i, \cdots c_P\}$ of a general attribute C in a two-level hierarchy, and $\{t_1, \cdots, t_i, \cdots t_p\}$ denotes the weights of the basic attributes, where $0 \leq t_i \leq 1$. There are M assessment grades.

Step 1. Convert the belief of each assessment level into belief. The conversion is shown in

Table 3: The antecedent attributes of security assessment

First level	Second level	Third level	Forth level
The Security Reliability of ICS(CS)	Security Event $(c_1^2)(t_1^2 = 0.5)$	The type of attacks $(c_{11}^2)c_{11}^2 = 0.2)$	Virus and trojan attack $(c_{111}^2)(t_{111}^2 = 0.3)$
			U2R attacks $(c_{112}^2)(c_{112}^2 = 0.2)$
			R2L attack $(c_{113}^2)(c_{113}^2 = 0.2)$
			Dos attack $(c_{114}^2)(c_{114}^2 = 0.3)$
		Continuous attack time $(c_{12}^2)(t_{12}^2 = 0.2)$	
		The frequcy of attacks $(c_{13}^2)(t_{13}^2 = 0.3)$	
		The Severity of attacks $(c_{14}^2)(t_{14}^2 = 0.3)$	
	Error Operatio $(c_2^2)(t_2^2 = 0.5)$	The Rate of events $(c_{21}^2)(t_{21}^2 = 0.5)$	
		The Severity of attacks $(c_{22}^2)(t_{22}^2 = 0.5)$	

Figure 5, where $Q_{i,j}$ represents the basic probabilistic set relative to the j-th assessment level $Q_{i,\Theta}$ is the rest of probabilistic set of the unassigned to any resultaccording to the i-th attribute, $\overline{Q_{i,\Theta}}$ denotes unallocated basic probability mass relative to the in significant degree of the i-th basic attribute, $Q_{i,\Theta}$ represents the unassigned basic probability mass with respect to the incompleteness of the i-th basic attribute.

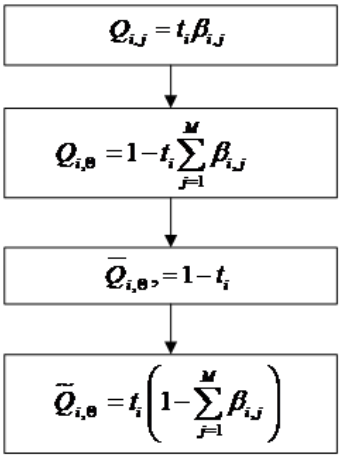


Figure 5: The conversion process

Step 2. Combine the first i-th attributes through the ER, and the detailed process can be describe as Figure 6.

where $Q_{I(i+1),j}$ represents the probability set of the j-th evaluation grade after the combi-

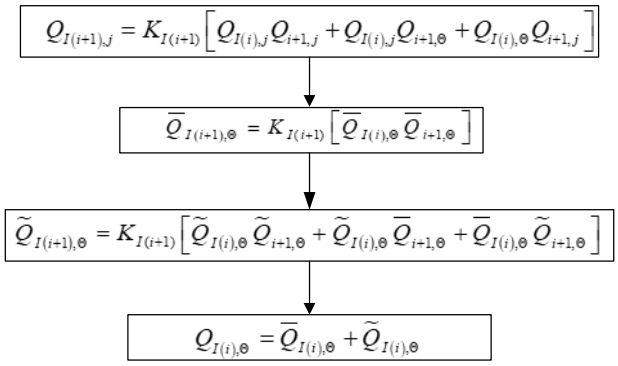


Figure 6: The combination process

nation of the first i-th basic attributes, and $K_{I(i+1)}$ can be calculated by:

$$K_{I(i+1)} = \frac{1}{1 - \sum_{k=1}^{N} \sum_{\substack{j=1 \\ j \neq k}}^{N} P_{I(i),k} P_{i+1,j}} \tag{6}$$

Step 3. Determine the belief of the j-th assessment level and the belief of the unassigned attributes by:

$$\widehat{\beta}_j = \frac{Q_{I(M),j}}{1 - \overline{Q}_{I(M),\Theta}} \quad (j = 1, 2, \cdots, N) \tag{7}$$

$$\widehat{\beta}_{\ominus} = \frac{\widetilde{Q}_{I(M),\Theta}}{1 - \overline{Q}_{I(M)\Theta}} \tag{8}$$

Compute the belief of the assessment results after determining all attributes:

$$\begin{aligned} &\beta_{i,j} = \tfrac{C_{i,j+1} - U(c_i)}{C_{i,j+1} - C_{i,j}} \left(C_{i,j+1} \leq U\left(c_i\right) \leq C_{i,j+1}\right) \\ &\beta_{i,j+1} = 1 - \beta_{i,j} \\ &\beta_{i,k} = 0 (k = 1, \cdots, N, k \neq j, j+1) \end{aligned} \tag{9}$$

where $U\left(c_i\right)$ is the value of attribute c_i and $c_{i,j}$ is the j-th reference value of c_i .

The ER rule can be described as an attribute fusion process. Based on the original data on the ICS, the various attributes should be mixed and then partially allocated to the third layer. Then, the assessment level of second layer attributes should be determined according to the third layer attributes, and that of first layer attributes according to the second layer attributes. Finally, the attributes of the three layers were integrated again by the ER rule, yielding to final ICS assessment result.

3.4 The reliability model based on BRB

The assessment results fused by ER rule should be graded according to the safety assessment result and fault assessment result, because the ICS reliability is an integration of fault and security. Then, the fused results will be taken as the input of the BRB model. As shown in Table 4, there are five reliability levels in FA and SC :

$$C^1 \in \{TW, B, G, E, TB\} \tag{10}$$

Table 4: Reliability level for ICS

Reliability Level	**The level of fault assessment (FA)**	**The level of security assessment (SC)**	**The level of Final result (FR)**
1(The worst)	TW	TW	TW
2(Bad)	B	B	B
3(Good)	G	G	G
4(Excellent)	E	E	E
5(The best)	TB	TB	TB

$$C^2 \in \{TW, B, G, E, TB\} \tag{11}$$

where C^1 is the reliability level of FA ; C^2 is the reliability level for of SC . The final result FR can be described as:

$$FR = (FR_1, FR_2, FR_3, FR_4, FR_5) = (\{TW, B, G, E, TB\}) \tag{12}$$

The reliability of R_k can be expressed as:

if $FA = C_k^1$ and $SC = C_k^2$, then $FR = \{(SC_1, t_{1,k}), \cdots, (SC_5, t_{5,k})\}$ with a rule weight θ_k and attribute weights ρ_{c^1} and ρ_{c^2} .

The BRB model have 25 rules in the initial case. The initial belief of the rules was determined by reliability assessment. If the fault assessment or security assessment are the worst (TW), it is impossible to guarantee the data accuracy in the ICS, i.e. the assessment result must the worst (TW). For example, the reliabilities of R_5, R_{10} , R_{15} , R_{20} , R_{25} can be expressed as:

$$\begin{array}{l} IfFAisTWandSCisTW, theFRis\{(TW,1),(B,0),(G,0),(E,0),(TB,0)\} \\ IfFAisBandSCisTW, theFRis\{(TW,1),(B,0),(G,0),(E,0),(TB,0)\} \\ IfFAisGandSCisTW, theFRis\{(TW,1),(B,0),(G,0),(E,0),(TB,0)\} \\ IfFAisEandSCisTW, theFRis\{(TW,1),(B,0),(G,0),(E,0),(TB,0)\} \\ IfFAisTBandSCisTW, theFRis\{(TW,1),(B,0),(G,0),(E,0),(TB,0)\} \end{array} \tag{13}$$

3.5 The reasoning process of BRB model

To obtain the stability of the ICS, the BRB model was derived by the ER method through the following steps:

Step 1. Calculate the matching degree of the antecedent attribute of the training sample to a rule:

$$a_i^k = \begin{cases} \dfrac{A_i^{l+1} - a_i(t)}{A_i^{l+1} - A_i^l} & k = l \\ \dfrac{a_i(t) - A_i^l}{A^{l+1} - A_i^l} & k = l+1 \\ 0 & k = 1, \cdots, K(k \neq l, l+1) \end{cases} \tag{14}$$

where a_i^k is the i-thantecedent rule; $a_i(t)$ is the i-th attribute in the data; A_i^l and A_i^{l+1} are the reference value of the l-th and ($l+1$) - thantecedent attributes, respectively; k represents the number of belief rules.

Step 2. Compute the weight of the k-th rule from the matching degree t_i^k :

$$t_k = \frac{\theta_k \prod_{i=1}^{M} \left(a_i^k\right)^{\overline{\delta}_i}}{\sum_{l=1}^{K} \theta_l \prod_{i=1}^{M} \left(a_i^l\right)^{\overline{\delta}_i}} \tag{15}$$

where $\overline{\rho}_i$ is the i-thantecedent attribute.

Step 3. Integrate the rules and generate belief of different outputs by the analytic synthesis algorithm of the ER which was described in Chapter 3.3.

Step 4. Generate the final output according to the belief of different outputs:

$$FA_{actual} = \sum_{n=1}^{N} C_n t_n \tag{16}$$

3.6 CMA-ES optimization of the BRB

The BRB was optimized under following three constraints: rule weight, attribute weight and rule. The rule weight constraint, reflecting the relative importance of the rule, can be expressed as:

$$0 \le \theta_k \le 1, \quad k = 1, 2, \ldots L \tag{17}$$

The attribute weight must be normalized to [0, 1]:

$$0 \le \rho_n \le 1, n = 1, \ldots, T_k \tag{18}$$

Similarly, the rule must be normalized to [0, 1]:

$$0 \le \rho_n \le 1, n = 1, \ldots, T_k \tag{19}$$

where $t_{i,k}$ is the belief of the k-th rule. The value of $t_{i,k}$=1 if the modified rule can be fully executed and smaller than one if otherwise:

$$\sum_{i=1}^{N} t_{i,k} \le 1, \quad k = 1, 2, \ldots, L \tag{20}$$

$$output_{estimated} = \sum_{n=1}^{N} p\left(SC_i\right) t_i \tag{21}$$

Next, the BRB parameters should be trained by the objective function to reduce the error between the assessed and actual outputs. Here, the error is measured by the mean square error (MSE):

$$\mathrm{MSE}\left(\theta_k, t_{i,k}, \rho_n\right) = \frac{1}{T} \sum_{i=1}^{T} (\text{ output }_{\text{ estimated }} - \text{ output }_{\text{ actual}})^2 \tag{22}$$

The objective function and constraints of the BRB parameter training can be expressed as:

$$\begin{aligned} &\min \mathrm{MSE}\left(\theta_k, t_{i,k}, \rho_n\right) \\ &0 \le \theta_k \le 1 \\ &0 \le \rho_n \le 1, n = 1, \ldots, T_k \\ &0 \le t_{i,k} \le 1, \quad i = 1, \ldots, N, \quad k = 1, 2, \ldots L \\ &\textstyle\sum_{i=1}^{N} t_{i,k} \le 1 \end{aligned} \tag{23}$$

The parameter training with the L-CMA-ES consists of four steps:

Step 1. Parameter initialization. The initial $mean^0$ is equal to the initial parameter Ω^0,and Ω is:

$$\Omega = \{\theta_1, \ldots, \theta_L, \beta_{1,\cdots}, \ldots, \beta_{N,}, \delta_1, \ldots, \delta_{T_k}\} \tag{24}$$

where the initial value of each parameter is given by the expert.

Step 2. Population sampling. Taking the central solution as the central expectation of the solution space, a normal distribution population is generated as:

$$\Omega_q^{g+1} \sim mean^g + \eta^g \mathbb{N}\left(0, \mathbf{C}^g\right)(q = 1, \cdots, n) \tag{25}$$

where Ω^0 is the first solution in the solution space; $\Omega_q^{g+1}$is the q-th solution in the solution space, g is the number of iterations; $mean$ is the mean distribution, i.e. the central expectation; η is the step size; C is the covariance matrix of the population.

Step 3. Selection and recombination. The candidate solution in the solution space are screened by leaky bucket mechanism. After the candidate solutions satisfying the constraints are obtained, the solutions of the child populations will be selected in the population according to the adaptive function. Then, the central expectation in the population moves towards the local optimal solution, and guides the evolution of the population. When the previous optimal solution population ε is obtained, the population's expectations can be updated by:

$$mean^{g+1} = \sum_{i=1}^{\varepsilon} \gamma_i \Omega_{i,\mu}^{g+1} \left(\sum_{i=1}^{\varepsilon} \gamma_i = 1\right) \tag{26}$$

where ε is the number of child populations; γ is individual weight (the total weight is 1); μ is the size of the parent population; $\Omega_{i:\mu}$ means the i-th candidate solution is obtained from the parent population μ in the $(g+1)$-th iteration according to the fitness value.

After recombination, the central region of the population will move towards the child populations, such that the candidate solution is more accurate than the parent population.

Step 4. Update of covariance matrix. In the next iteration, the optimal solution needs to be found based on the covariance matrix. During the iteration process, the transformation of the covariance matrix varies with the length and orientation of the long axis of the elliptical distribution of the population. The change in orientation reflects the trend and direction of evolution, while the change in length represents the scope of the population search. The covariance matrix should be updated by:

$$\mathbf{C}^{g+1} = \left(1 - a_1 - a_\varepsilon\right)\mathbf{C}^g + a_1 p^{g+1}\left(p^{g+1}\right)^T + a_\varepsilon \sum_{i=1}^{\varepsilon} \gamma_i \left(\frac{\left(\Omega_{i,\mu}^{g+1} - mean^g\right)}{\eta^g}\right)\left(\frac{\left(\Omega_{i,\mu}^{g+1} - mean^g\right)}{\eta^g}\right)^T \tag{27}$$

where a_i and a_ε is the total learning rate; p is the evolution path (initial value=0). Then, the evolution path should be updated as:

$$p^{g+1} = \left(1 - a_p\right)p^g + \sqrt{a_p\left(2 - a_p\right)\left(\sum_{i=1}^{s} \gamma_i^2\right)^{-1}} \frac{mean^{g+1} - mean^g}{\eta^g} \tag{28}$$

where $a_p \leq 1$ is the retrospective period of the evolution path. The step η should be updated as:

$$\eta^{g+1} = \eta^g \exp\left(\frac{a_\eta}{d_\eta}\left(\frac{\left|\left|p_\eta^{g+1}\right|\right|}{\mathrm{E}||\mathbb{N}(0, \mathbf{I})||} - 1\right)\right) \tag{29}$$

where d_η is the damping coefficient; $\mathrm{E}||\mathrm{N}(0, \mathrm{I})||$ is the expectation of the Euclidean paradigm $||N(0, \mathbf{I})||$; I is the vector of the unit matrix; $a_\eta$is the look back window; p_η is the conjugate evolution path (initial value=0). The conjugate evolution path should be updated as:

$$p_{\eta}^{g+1} = (1 - a_{\eta})\, p_{\eta}^{g} + \sqrt{a_{\eta}\,(2 - a_{\eta}) \left(\sum_{i=1}^{\varepsilon} \gamma_i^2 \right)^{-1}}\, \mathbf{C}^{(g)-\frac{1}{2}} \frac{m^{g+1} - m^{g}}{\eta^{g}} \tag{30}$$

The above formula(27-30) should be executed repeatedly until reaching the accuracy requirement. Then, the optimal solution should be outputted, and serve as the model inputs after training.

4 Case study

4.1 The assessment grades of the attributes

Since the ICS reliability is affected by external and internal factors, this section designs simulation experiments involving both internal faults and external security incidents. The experiments were carried out in the actual industrial control environment. The faults and incidents were selected and rated empirically by experts. Some of them are quantitative information, and some are qualitative. Both internal faults and external security incidents were divided into four levels according to the actual situation (Table 5). The dataset was derived from the log events of a Chinese tobacco factory, which were recorded over 100 days by a PLC-controlled device in SCADA system.

Table 5: The levels of internal faults

	Worst	**Bad**	**Good**	**Best**
c_{11}^{1}(times/day)	15	10	5	0
c_{12}^{1}(hours)	3	12	21	30
c_{13}^{1}(minutes)	60	40	20	10
c_{14}^{1}(minutes)	(minutes)	60	40	20
c_{15}^{1}(hours)	3	12	21	30
c_{16}^{1}	Given by Experts			
c_{17}^{1}	Given by Experts			
c_{21}^{1}(hour)	3	12	21	30
c_{22}^{1}(minutes)	1	40min	20min	10
c_{23}^{1}	Given by Experts			
c_{24}^{1}(hour)	3	12	21	30
c_{31}^{1}(times/ day)	5	3	1	0
c_{32}^{1}(%)	10	5	3	0%
c_{33}^{1}(hour)	10	20	30	40
c_{34}^{1}	Given by Experts			
c_{35}^{1}(seconds)	1s	0.3s	0.1s	0.01s

4.2 Internal and external reliabilities based on ER rule

The original data were fused by the ER algorithm, using the attributes and reference levels in Table 5-6. The fusion follows the process detailed in Section 3. According to the ER algorithm, the author obtained quantitative observation data for 30 days, which reflect the internal reliability

of the ICS. The daily observation data are shown in Figure 7. It can be seen from Figure 7 that the internal reliability of the ICS fluctuated violently, indicating a high frequency of system faults, while the external reliability changed less significantly, which reflects the moderate frequency of external attacks.

Table 6: The levels of external security incidents

	Worst	**Bad**	**Good**	**Best**
c_{111}^2	Given by Experts			
c_{112}^2	Given by Experts			
c_{113}^2	Given by Experts			
c_{114}^2	Given by Experts			
c_{12}^2(minutes)	20	10	5	0
c_{13}^2(times/day)	5	3	1	0
c_{14}^2	Given by Experts			
c_{21}^2(times/day)	5	3	1	0
c_{22}^2	Given by Experts			

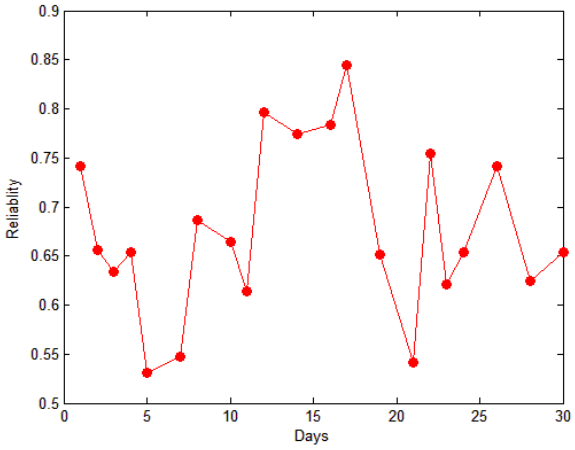


Figure 7: Internal reliabilities for 30 days

4.3 Model construction

The proposed method (BRB) was contrasted with two prediction models, i.e. back propagation neural network (BP) and Markov prediction model (MM). For consistency, the initial parameters of all three methods were trained by the CMA-ES algorithm. The BP is a popular data-based prediction model. This model takes Gaussian functions as training neurons, and adopts radial basis function kernels in the middle layer for nonlinear transform of input parameters. Compared with traditional neural networks, the BP enjoys a small scale and fast operation, thanks to the limited number of intermediate layers.

The MM is a typical semi-quantitative assessment model based on bayesian decision theory. It bears high resemblance to the proposed BRB model. The high accuracy and efficiency have earned it immense popularity.

The initial weights and levels of belief rules are listed in Table 7. The initial parameters of the two contrastive algorithms are given in Table 8. The input parameters of the CMA-ES

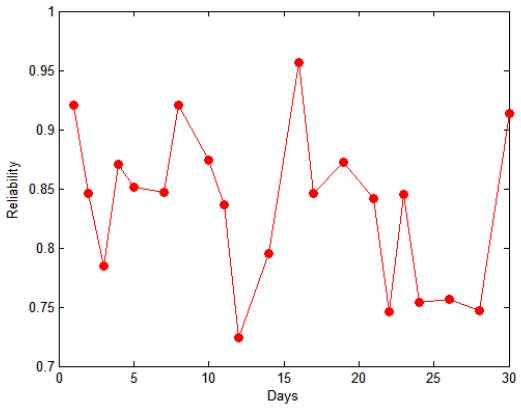


Figure 8: External reliabilities for 30 days

Table 7: Initial weights and levels of belief rules

Rule	Rule Weight	SC(t)	SC1, SC2, SC3, SC4, SC5=TW, B, G, E, TB
1	1	TW	$(SC_1,1),(SC_2,0),(SC_3,0),(SC_4,0),(SC_5,0),(SC,0)$
2	1	B	$(SC_1,0),(SC_2,1),(SC_3,0),(SC_4,0),(SC_5,0),(S,0)$
3	1	G	$(SC_1,0),(SC_2,0),(SC_3,1),(SC_4,0),(SC_5,0),(SC,0)$
4	1	E	$(SC_1,0),(SC_2,0),(SC_3,0),(SC_4,1),(SC_5,0),(SC,0)$
5	1	TB	$(SC_1,0),(SC_2,0),(SC_3,0),(SC_4,0),(SC_5,1),(S,0)$

Table 8: The initial parameters of the two contrastive methods

Comparison	Initial parameters
MM	Initial probability vector = [0.2,0.2,0.2,0.2,0.2];
	Initial probability transition matrix = [1,0,0,0,0;0,1,0,0,0;0,0,1,0,0;0,0,0,1,0;0,0,0,0,1];
BP	Input neuron: 3;Output neuron: 1;
	Sliding window size: 3

Table 9: The input parameters of the CMA-ES algorithm

The semantic	Initial parameters
value	$m^0 = O^0$; $\sigma^0 = 0.5; \lambda = 13$; $\tau = 6$; $a_1 = 0.0031$; $a_\tau = 0.0066$; $p_\psi^0 = 0$; $a_\psi = 0.147$; $p_\sigma^0 = 0$; $a_\sigma = 0.1813$; $d_\sigma = 1.1813$; $e = 0.6$; Loop = 100;

algorithm are displayed in Table 9.

4.4 Simulation and result analysis

Ten rounds of validations were performed to verify the effectiveness of our method. The exact range of the result was determined through interval estimation.The 100 reliability values

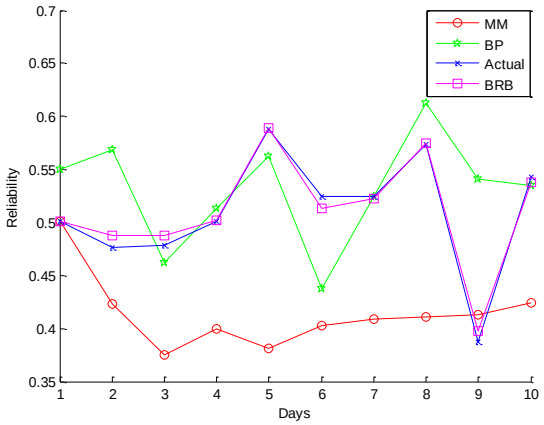


Figure 9: The results of round 1

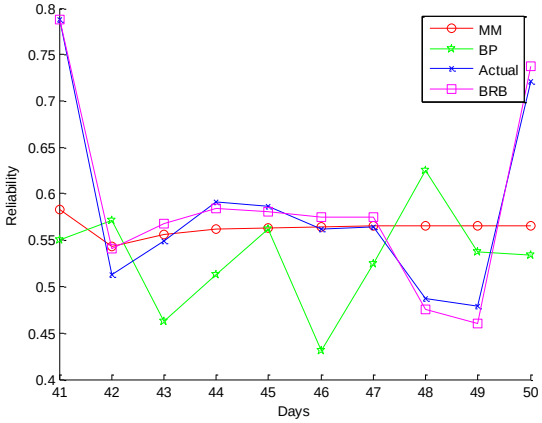


Figure 10: The results of round 5

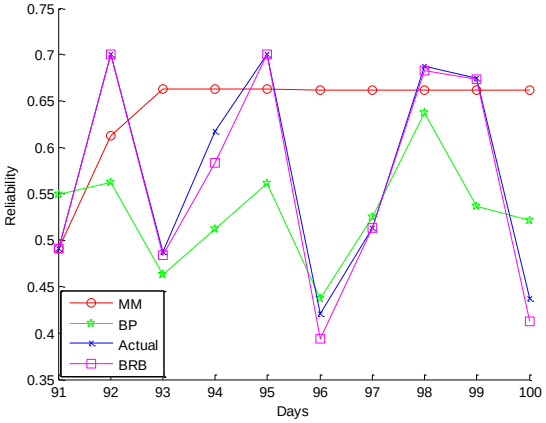


Figure 11: The results of round 10

were divided into 10 groups. The first 9 groups were adopted as the training data, and the rest as the test data. The results of the first, fifth and tenth round of training are presented in Figure 9, 10 and 11, respectively.

Table 10: The MSEs after round 10

Rounds	**1**	**2**	**3**	**4**	**5**
BRB	$4.676e^{-5}$	$3.097e^{-5}$	$1.345e^{-4}$	$0.987e^{-4}$	$1.362e^{-4}$
MM	$0.963e^{-2}$	$6.698e^{-3}$	$2.322e^{-2}$	$6.342e^{-2}$	$3.987e^{-2}$
BP	$2.547e^{-3}$	$3.147e^{-3}$	$6.357e^{-3}$	$7.214e^{-3}$	$1.365e^{-2}$
Rounds	**6**	**7**	**8**	**9**	**10**
BRB	$7.321e^{-4}$	$8.541e^{-5}$	$1.247e^{-4}$	$6.218e^{-5}$	$1.361e^{-4}$
MM	$0.897e^{-2}$	$3.657e^{-3}$	$3.465e^{-2}$	$7.645e^{-2}$	$1.644e^{-2}$
BP	$1.247e^{-2}$	$4.365e^{-3}$	$4.514e^{-3}$	$5.987e^{-3}$	$7.365e^{-3}$

Table 10 lists the MSEs between the predicted results and the actual results of the three methods after round 10.

Table 11 provides the overall MSEs and the interval estimation results of the three methods after round 10. The interval estimation results refer to the mean estimation interval at the confidence level of 95%.

Table 11: The results of overall MSEs and the interval estimation after round 10

	MSE	**Interval estimation (95%)**
BRB	2.36 e^{-4}	[2.12 e^{-4}, 3.87 e^{-4}]
MM	2.16 e^{-4}	[1.37 e^{-2}, 3.98 e^{-2}]
BP	6.37 e^{-3}	[3.780 e^{-3}, 8.98 e^{-3}]

The above results show that the proposed BRB model, the MM and the BP respectively controlled the MSE on the order of 1,000th, 100th and 1,000th. Thus, our model reduced the error by 100 times from the level of the MM. Moreover, our model achieved a small and acceptable error range. Thus, the proposed method is reliable and efficient, in addition to its good prediction accuracy.

5 Conclusions

Considering the internal and external influencing factors of ICS reliability, this paper puts forward a new ICS reliability assessment model based on the BRB. The ICS reliability assessment was divided into internal fault assessment and external security assessment, making the assessment more objective. The BRB can process semi-quantitative information in our samples, which contain both quantitative data and qualitative knowledge. In this way, the model can be trained well with sufficient samples. The initial training parameters were optimized by the CMA-ES algorithm, aiming to improve the accuracy of model inputs. Finally, the proposed model was compared with two other popular prediction methods through case study. The results show that the proposed method is reliable, efficient and accurate, laying a solid basis for reliability assessment of complex ICSs.

Funding

This work was supported in part by the Scientific Research Starting Foundation for Returned Overseas of Heilongjiang Province under Grants LC2018030.

Author contributions. Conflict of interest

The authors contributed equally to this work. The authors declare no conflict of interest.

Bibliography

[1] Alambeigi, F.; Wang, Z.; Hegeman, R. (2019). Autonomous data-driven manipulation of unknown anisotropic deformable tissues using unmodelled continuum manipulators, *IEEE Robotics and Automation Letters*, 4(2), 254-261, 2019.

[2] Brereton, R. G. (2016). Hotelling's T squared distribution, its relationship to the F distribution and its use in multivariate space, *Journal of Chemometrics*, 30(1), 18-21, 2016.

[3] Chen, Q.; Abercrombie, R. K.; Sheldon, F. T. (2015). Risk assessment for industrial control systems quantifying availability using mean failure cost (MFC), *Journal of Artificial Intelligence and Soft Computing Research*, 5(3), 205-220, 2015.

[4] Franco, I. C.; Schmitz, J. E.; Costa, T. V. (2017). Development of a predictive control based on Takagi-Sugeno model applied in a nonlinear system of industrial refrigeration, *Chemical Engineering Communications*, 204(1), 39-54, 2017.

[5] Goedhart, R.; Schoonhoven, M.; Ronald, J. M. M. (2016). Correction factors for Shewhart and control charts to achieve desired unconditional ARL, *International Journal of Production Research*, 54(24), 7464-7479, 2016.

[6] He, W.; Hu, G. Y.; Zhou, Z. J. (2018). A new hierarchical belief-rule-based method for reliability evaluation of wireless sensor network, *Microelectronics Reliability*, 87, 33-51, 2018.

[7] Hu, G. Y.; Zhou, Z. J.; Zhang B. C. (2016). A method for predicting the network security situation based on hidden BRB model and revised CMA-ES algorithm, *Applied Soft Computing*, 48(C), 404-418, 2016.

[8] Jin, L. J.; Peng, C. Y.; Jiang, T. (2017). System-level electric field exposure assessment by the fault tree analysis, *IEEE Transactions on Electromagnetic Compatibility*, 59(4), 1095-1102, 2017.

[9] Kriaa, S.; Pietre, L.; Bouissou, M. (2015). A survey of approaches combining safety and security for industrial control systems, *Reliability Engineering & System Safety*, 139, 156-178, 2015.

[10] Lee, Y. S.; Kim, D. J.; Kim, J. O. (2011). New FMECA methodology using structural importance and fuzzy theory, *IEEE Transactions on Power Systems*, 26(4), 2364-2370, 2011.

[11] Liu, Z.; Liu, T.; Han, J. (2017). Signal model-based fault coding for diagnostics and prognostics of analog electronic circuits, *IEEE Transactions on Industrial Electronics*, PP(99), 1-1, 2017.

[12] Luo, Z. Y.; Wang, P.; You, B. (2016). Serial reduction optimization research of complex product workflow's accuracy under the time constraint, *Advances in Mechanical Engineering*, 8(10), 1-9, 2016.

[13] Luo, Z. Y.; You, B.; Liu, J. H. (2016). Research of the intrusion tolerance state transition system based on semi-markov, *Transactions of Beijing Institute of Technology*, 36(7), 712-717, 2016.

[14] Luo, Z. Y.; You, B.; Xu, Z. B. (2014). Automatic recognition model of intrusive intention based on three layers attack graph, *Jilin Daxue Xuebao (Gongxueban)/Journal of Jilin University (Engineering and Technology Edition)*, 44(5), 1392-1397, 2014.

[15] Navarro, A. D.; Yip, H. M.; Wang, Z. (2016). Automatic 3-d manipulation of soft objects by robotic arms with an adaptive deformation model, *IEEE Transactions on Robotics*, 32(2), 429-441, 2016.

[16] Pacella, M. (2018). Unsupervised classification of multichannel profile data using PCA: An application to an emission control system, *Computers & Industrial Engineering*, 122, 161-169, 2018.

[17] Rusmini, P.; Crippa, V.; Cristofani, R. (2016). The role of the protein quality control system in SBMA, *Journal of Molecular Neuroscience*, 58(3), 348-364, 2016.

[18] Sang, W.; Livne, E. (2016). Probabilistic aeroservoelastic reliability assessment considering control system component uncertainty, *Aiaa Journal*, 54(8), 2507-2520, 2016.

[19] Thomas, M. C.; Zhu, W.; Romagnoli, J. A. (2017). Data mining and clustering in chemical process databases for monitoring and knowledge discovery, *Journal of Process Control*, S095915241730032X, 2017.

[20] Wang, Z.; Li, P.; Navarro, A. D. (2015). Design and control of a novel multi-state compliant safe joint for robotic surgery, *IEEE International Conference on Robotics and Automation (ICRA)*, 1023-1028, 2015.

[21] Wang, Z.; Yip, H. M.; Navarro, A. D. (2016). Design of a novel compliant safe robot joint with multiple working states, *IEEE/ASME Transactions on Mechatronics*, 21(2), 1193-1198, 2016.

[22] Zhou, Z. J.; Hu, G. Y.; Zhou, Z. J. (2017). A Model for hidden behavior prediction of complex systems based on belief rule base and power set, *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, PP(99), 1-7, 2017.

INTERNATIONAL JOURNAL OF COMPUTERS COMMUNICATIONS & CONTROL
ISSN 1841-9836, e-ISSN 1841-9844, 14(3), 437-452, June 2019.

Performer Selection in Human Reliability Analysis: D numbers Approach

J. Zhao, Y. Deng

Jie Zhao
Institute of Fundamental and Frontier Science
University of Electronic Science and Technology of China, Chengdu

Yong Deng*
Institute of Fundamental and Frontier Science
University of Electronic Science and Technology of China, Chengdu
*Corresponding author: dengentropy@uestc.edu.cn

Abstract: Dependence assessment among human errors in human reliability analysis (HRA) is an significant issue. Many previous works discussed the factors influencing the dependence level but failed to discuss how these factors like "similarity of performers" determine the final result. In this paper, the influence of performers on HRA is focused, in addition, a new way of D numbers which is usually used to handle with the multiple criteria decision making (MCDM) problems is introduced as well to determine the optimal performer. Experimental result demonstrates the validity of proposed methods in choosing the best performers with lowest the conditional human error probability (CHEP) under the same circumstance.
Keywords: Human reliability analysis (HRA), D numbers, D-S evidence theory, multiple criteria decision making (MCDM), nuclear power plant.

1 Introduction

Human error is an important factor to be considered in the design and risk assessment of large complex systems [71], such as nuclear power plant operation [1, 28, 66], air traffic control [27, 44], grounding of oil tankers [37] and IIR filters [40]. Dependence analysis within human reliability analysis (HRA) refers to evaluating the influence of the failure of the operator to perform one task on the failure probabilities of subsequent tasks [51]. Dependence between two sequent tasks within human reliability analysis means if the preceding task fails, the failure probability of the following task is higher than the success [72]. Therefore, an appropriate assessment of dependence is crucial to avoid underestimation of the risk. In the existing methods, the result of dependence assessment is a Conditional Human Error Probability (CHEP), given failure on the preceding task [5, 72].

The most widely used method is the technique for human error rate prediction (THERP) dependence method [26, 50]. THERP introduces five levels of dependence corresponding to df-ferent values of CHEPs and it suggests some of the factors that may influence the dependence level. The THERP model refers to three main factors: "Closeness in Time" (CT), "Task Relatedness" (TR) and "Similarity of Performers" (SP). And the assessment requires considerable expert judgment on identifying which factor is important and on how these factors influence the dependence level. As a result, a highly subjective process that may be insufficient traceability and reproducibility would be inevitable.

To overcome this limitation, Decision Trees (DTs) method [2,17] and Fuzzy Expert System (FES) [38, 39, 72] have been introduced to extend the THERP dependence model. However, the influence of performer is partially ignored to some degree in the existing methods, which is critical factor determining the dependence level. In addition, it is inevitable to deal with uncertainty in the modeling the influence of performer. As a result, it is necessary to develop a new HRA

math model. The purpose of this paper is to present a model and discuss how the differences of performers affect the final result as well as conditional human error probability (CHEP). The proposed method is based on D numbers [10] which is efficient to handle uncertainty multiple criteria decision making problem (MCDM) [9, 55, 56].

The paper is organized as follows. In Section 2, some preliminaries are briefly introduced, including HRA, D-S evidence theory and D numbers. In Section 3, the proposed methodology in HRA is detailed. In Section 4, an application of the proposed HRA method in post initiator HFEs of a nuclear power plant is used to illustrate the effectiveness of the presented method. Conclusion is given in Section 5.

2 Preliminaries

In this section, some preliminaries including human reliability analysis (HRA), evidence theory and D numbers, are briefly introduced.

Definition 1. Assume that task T_B is subsequent to task T_A, and B and A are the corresponding failure events. P_A and P_B are basic probability of failure of task T_A and T_B, respectively, then the conditional human error probability (CHEP) of B given is defined as follows: [47]

$$P_{XD}(B|A) = (1 + K \times P_B)/(K + 1) \tag{1}$$

where K=0, 1, 6, 9, ∞, for dependence levels CD, HD, MD, LD and ZD, where XD=CD, HD, MD, LD and ZD, respectively. It reflects the less dependent the two tasks are, the lower the failure of probability is.

2.1 D-S evidence theory

It's necessary to deal with uncertainty [34, 35, 63, 65] and many math models are presented such as fuzzy set [12,41,53,64], Z-numbers [24,25,59], belief function [20,60,61] and bio-inspired model [62]. Due to its efficiency to model uncertainty, evidence theory is widely used in decision making [3, 14], pattern recognition [4, 6, 19] evidential reasoning [29, 67–69] and information fusion [46, 58]. Some basic definitions of D-S theory are briefly introduced [7, 42]:

Definition 2. A set of hypotheses Θ is the exhaustive hypotheses of variable and the elements are mutually exclusive in Θ. Then Θ is called the frame of discernment, defined as follows [7,42]:

$$\Theta = \{H_1, H_2, \cdots, H_i, \cdots, H_N\} \tag{2}$$

The power set of Θ is denoted by 2^Θ, and

$$2^\Theta = \{\emptyset, \{H_1\}, \cdots, \{H_N\}, \{H_1, H_2\}, \cdots, \{H_1, H_2, \cdots, H_i\}, \cdots, \Theta\} \tag{3}$$

where $\emptyset$ is an empty set.

Definition 3. A BPA function m is a mapping of 2^Θ to a probability interval $[0, 1]$, formally defined by [7, 42]:

$$m: \quad 2^\Theta \to [0, 1] \tag{4}$$

which satisfies the following conditions:

$$m(\emptyset) = 0 \quad \sum_{A \in 2^\Theta} m(A) = 1 \quad 0 \le m(A) \le 1 \quad A \in 2^\Theta \tag{5}$$

The mass $m(A)$ represents how strongly the evidence supports A with the efficiency to model uncertainty [9]. Conflicting management with Dempster rule is still an open issue [52]. In addition, many operations on mass function have been proposed such correlation [21], distance [11], entropy measure [31,36], divergence measure [13,45] and negation [15,16].

For the same evidence, the different BPAs come from the different evidence resources. The Dempster's combination rule can be used to obtain the combined evidence [7]:

$$\begin{cases} m(\emptyset) = 0 \\ m(A) = \frac{\sum\limits_{B \bigcap C = A} m_1(B) m_2(C)}{1-K} \end{cases} \tag{6}$$

where $K = \sum\limits_{B \bigcap C = \emptyset} m_1(B) m_2(C)$. It should be mentioned that many methods are presented to deal with the open issues of evidence theory.

Definition 4. Let m be a BPA on Θ. Its associated pignistic probability function Bet P_m: $\Theta \to [0,1]$ is defined as:

$$Betp_m(\omega) = \sum_{A \in \Theta, \omega \in A} \frac{1}{|A|} \frac{m_A}{1 - m_(\emptyset)}, m(\emptyset) \neq 1, \tag{7}$$

where $|A|$ is the cardinality of subset A.

2.2 D numbers

The real application exists the uncertainty [22,57]. To address the Multiple criteria decision making problem (MCDM) issue [23], a new mathematical tool called D numbers [10] to represent uncertain information is proposed. The D numbers overcome several limitations the D-S theory involved.

Definition 5. Let Ω be a finite nonempty set, a D number is a mapping that D: $\Omega \to$ [0,1],with

$$\sum_{A \in \Omega} D(A) \leq 1 \quad and \quad D(\emptyset) = 0. \tag{8}$$

where $\emptyset$ is an empty set and A is a subset of Ω [10].

Besides the empty set is not necessary set ZERO [48, 49], it should be pointed out that different from D-S theory, the elements in set ω do not require mutually exclusive and the completeness constraint is not necessary in D numbers. If $\sum\limits_{A \in \Omega} D(A) \leq 1$, the information is said to be complete; Otherwise, the information is assumed to be incomplete.

For a discrete set Ω=$\{b_1$,b_2,$\cdots$,b_i,b_j,$\cdots$,$b_n\}$, where $b_i \in$R and $b_i \neq b_j$, when $i \neq j$,a special form of D numbers can be expressed by

$$\begin{aligned} D(b_1) &= v_1 \\ D(b_2) &= v_2 \\ &... \\ D(b_i) &= v_i \\ D(b_j) &= v_j \\ &... \\ D(b_n) &= v_n \end{aligned} \tag{9}$$

simply denoted as D=$\{(b_1, v_1),(b_2, v_2),\cdots,(b_i, v_i),(b_j, v_j),\cdots,(b_n, v_n)\}$, where $v_i > 0$ and $\sum_{i=1}^{n} v_i \leq 1$. If $\sum_{i=1}^{n} v_i = 1$, the information is said to be complete; If $\sum_{i=1}^{n} v_i < 1$, the information is said to be incomplete. It is effective and convenient that using the form of D numbers to express the uncertain information in the real world.

Definition 6. Let $D_1 = \{(b_1^1, v_1^1), \cdots, (b_i^1, v_i^1), \cdots, (b_n^1, v_n^1)\}, D_2 = \{(b_1^2, v_1^2), \cdots, (b_j^2, v_j^2), \cdots, (b_m^2, v_m^2)\}$ be two D numbers, the combination of D_1 and D_2,indicated by $D = D_1 \bigoplus D_2$,is defined by

$$D(b) = v \tag{10}$$

with

$$b = \frac{b_i^1 + b_j^2}{2} \tag{11}$$

$$v = \frac{v_i^1 + v_j^2}{2}/C \tag{12}$$

$$C = \begin{cases} \sum_{j=1}^{m}\sum_{i=1}^{n}(\frac{v_i^1+v_j^2}{2}), & \sum_{i=1}^{n} v_i^1 = 1 \quad ; \ and \quad \sum_{j=1}^{m} v_j^2 = 1; \\ \sum_{j=1}^{m}\sum_{i=1}^{n}(\frac{v_i^1+v_j^2}{2}) + \sum_{j=1}^{m}(\frac{v_c^1+v_j^2}{2}), & \sum_{i=1}^{n} v_i^1 < 1 \quad and \quad \sum_{j=1}^{m} v_j^2 = 1; \\ \sum_{j=1}^{m}\sum_{i=1}^{n}(\frac{v_i^1+v_j^2}{2}) + \sum_{j=1}^{m}(\frac{v_i^1+v_c^2}{2}), & \sum_{i=1}^{n} v_i^1 = 1 \quad and \quad \sum_{j=1}^{m} v_j^2 < 1; \\ \sum_{j=1}^{m}\sum_{i=1}^{n}(\frac{v_i^1+v_j^2}{2}) + \sum_{j=1}^{m}(\frac{v_c^1+v_j^2}{2}) & \\ + \sum_{j=1}^{m}(\frac{v_i^1+v_c^2}{2}) + \frac{v_c^1+v_c^2}{2}, & \sum_{i=1}^{n} v_i^1 < 1 \quad and \quad \sum_{j=1}^{m} v_j^2 < 1; \end{cases} \tag{13}$$

where $v_c^1 = 1 - \sum_{i=1}^{n} v_i^1$ and $v_c^2 = 1 - \sum_{j=1}^{n} v_j^2$

In the meanwhile, an aggregation operator is proposed on this special D number, it is defined as below.

Definition 7. Let$D = \{(b_1, v_1), (b_2, v_2), \cdots, (b_i, v_i), \cdots, (b_n, v_n)\}$be a D number, the integration representation of D number is defined as:

$$I(D) = \sum_{i=1}^{n} b_i v_i \tag{14}$$

3 The proposed method

The motivation behind the development of the proposed method is to select best performers which is the key factor in HRA among alternatives using D numbers proposed by Deng [10]. The influence of performers should be paid more attention. After figuring out the best performers

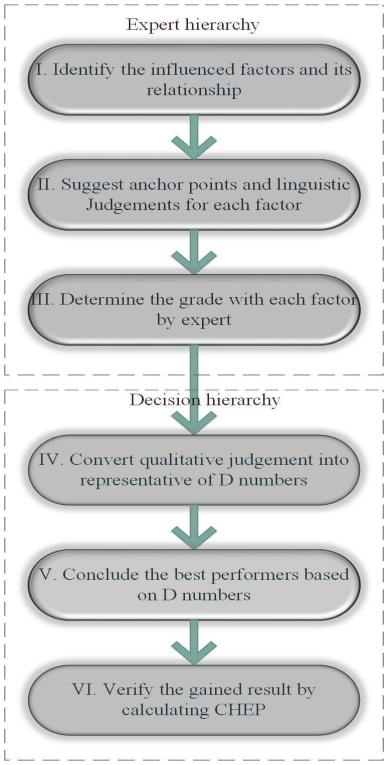


Figure 1: Framework of the proposed method

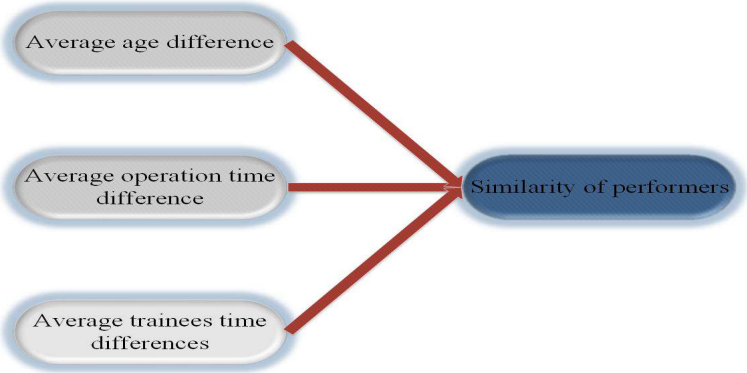


Figure 2: The factor affecting the performers

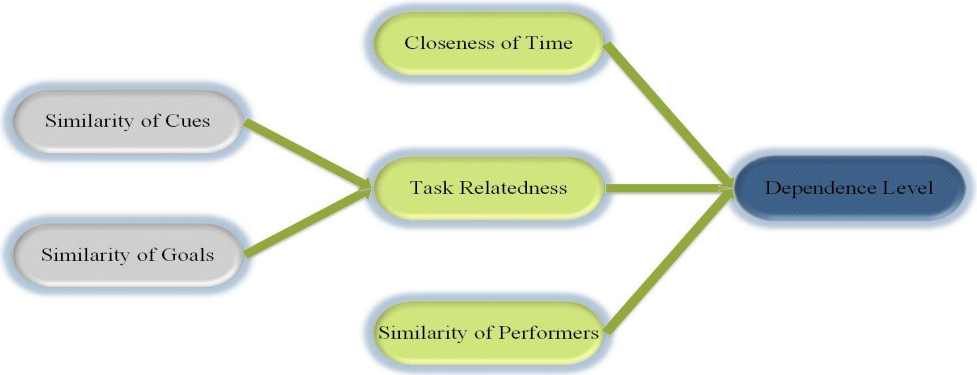


Figure 3: Functional relationships among the input factors of the dependence level

based on D numbers, the result through calculating the CHEP will be verified. The flow chart is shown in Fig. 1.

Step 1. Identify the influenced factors and the functional relationship among them: The first step is to determine the factors that may have influence on the dependence between two human failure events. For example, the THERP [72] model refer to five main factors: "closeness in time", "task relatedness", "similarity of performers", "similarity of cues" and "similarity of goals" .And as for the performers, the influential factors the similarity of performers could be divided into three: "the average age among performers", "the average trainees time among performers" and "average operation time on the same task among performers". The relationship is shown in Fig. 2 and Fig. 3.

Step 2. Suggest anchor points and linguistic judgements for each factor: anchor points and linguistic judgements are provided by domain experts in advance as guidance for the HRA analyst's judgement of the input factors. Furthermore, due to the frequent ignorance of performers, the influential factors "the average age among performers", "the average trainees time among performers" and "average operation time on the same task among performers" are proposed. Their anchor points are shown in Tables (1-3). And a set of anchor points and linguistic judgements for the input factor SP is presented in Table 4, with the associated levels of dependence of performers. Dependence level of an input factors indicates the dependence level between two tasks with respect to this factor. For example, dependence level "ZD" in Table 4 means that the dependence level between two tasks is zero dependence with respect to the factor "similarity of performers".

"Average operation time difference" anchor points	Linguistic judgement	Dep. level
Less than 30 min	The tasks are accomplished by the same individual in operation time	CD
More than 30 min but less than 1h	High level of performer similarity in operation time exists	HD
More than 1h but less than 1.5h	The level of performer similarity in operation time is medium	MD
More than 1.5h but less than 2h	A low level of performers similarity in operation time exists	HD
More than 2h	NO similarity of performers in operation time is present between tasks	ZD

Table 1: Anchor points for input factor "average operation time difference"

"Average age difference "anchor points	Linguistic judgement	Dep. level
less than 1 year	The age of the two performers are basically the same	CD
more than 1 year but less than 3 years	The age similarity of two performers are almost the same	HD
more than 3 years but less than 5years	The age level of performers similarity is medium	MD
more than 5 years but less than 7years	The difference of the two performers are quite different in age	HD
more than 7 years	The tasks are accomplished by the completely different individuals	ZD

Table 2: Anchor points for input factor "average age difference"

"Average trainees time difference" anchor points	Linguistic judgement	Dep. level
less than 1 year	The operating technology between two performers is basically same	CD
more than1 year but less than 1.5 year	The technology difference between two performers are slight	HD
more than 1.5 years but less than 2 years	The level of operating technology between two performers is medium	MD
more than 2 years but less than 2.5 years	The operating technology between two performers is basically different	HD
more than 2.5 years	The technology of the two teams are different	ZD

Table 3: Anchor points for input factor "average trainees time difference"

Step 3. Determine the grade with each factor by expert: It is supposed that the expert is totally authoritative. According to the collected data, the expert is asked to give their qualified judgement on each factor. The factor need to be judged by the expert are: the average age among performers, the average trainees time among performers, average operation time on the same task among performers, "closeness in time", "task relatedness". Due to our purpose is to verify best performers, the factors "closeness in time", "task relatedness" are set as constant.

Anchor points	Linguistic judgement	Dep. level
TSC vs control shift room	NO similarity of performers is present between tasks	ZD
Different team	A low level of performer similarity exist	LD
Different individuals with same qualification	The level of performer similarity is medium	MD
Same team	High level of performer similarity is present between tasks	HD
Same person	The tasks are accomplished by the same individual	CD

Table 4: Judgement on factors given by experts after transformation

Step 4. Convert qualitative judgement into representation of D numbers: In the previous work in the HRA, the lack of objectivity is always a severe problem. The highlight of this paper is the representation of the expert's judgement, by which could effectively represent the judgement in an reasonable way as well as offer a new way to deal with Multiple criteria decision making problem (MCDM). The transformation from the judgement into the D numbers is shown in Eq. (9).
Take the "the average age among performers" of performer A as an example:
$D^A_{age} = \{(HD, 0.6), (CD, 0.4)\}$.

Step 5. Conclude the best performers based on D numbers: given the existed judgement, we can calculate the utility of each factor using Eqs. (8-13). Firstly aggregating the three influential factor one by one, then the utility of performer is obtained.
Take the performer A as an example:
$D^A_{age} = \{(HD, 0.6), (CD, 0.4)\}$
$D^A_{ope} = \{(HD, 0.3), (CD, 0.7)\}$
$D^A_{time} = \{(CD, 1.0)\}$

To simplify the representation and calculation, it translates the qualitative judgement into numerical form: the dependence level CD, HD, MD, LD, ZD are corresponding to 5, 4, 3, 2, 1 respectively. The rest of the paper would follow this simplification.
$D^A_{age} = \{(4, 0.6), (5, 0.4)\}$
$D^A_{ope} = \{(4, 0.3), (5, 0.7)\}$
$D^A_{time} = \{(5, 1.0)\}$

According to Eqs. (9-12), the integrated D numbers by fusing above three D numbers is obtained:
$D_{performerA} = D^A_{age} \bigoplus D^A_{ope} \bigoplus D^A_{time}$
$D_{performerA} = \{(4.5, \frac{4.9}{16}), (4.75, \frac{6}{16}), (5, \frac{5.1}{16})\}$

Then using Eq. (14), it can conclude:
$I(D_{performerA}) = 5 \times \frac{5.1}{16} + 4.75 \times \frac{6}{16} + 4.5 \times \frac{4.9}{16} = 4.7531$

In the transformation from expert's judgements into D numbers, the more dependent the influential factors of SP are, the higher value the D numbers would be. Therefore, it is could be said that by comparing the I(D) of each factor, the smaller number of I(D) is, the better performer is. Then the best performer is obtained.

In this paper, a transformation form D number to BPA is defined. Assuming there are several D numbers of one individual, the translation from D numbers into BPA is defined as:

$$\begin{aligned} D^1_i &= \{(b_1, v_1), (b_2, v_2)\} \\ D^1_j &= \{(b_1, v_3), (b_2, v_4)\} \\ D^1_k &= \{(b_1, v_5), (b_2, v_6)\} \end{aligned} \tag{15}$$

the BPA is obtained:

$$m(b_1) = \frac{v_1 + v_3 + v_5}{n}, m(b_2) = \frac{v_2 + v_4 + v_6}{n} \tag{16}$$

where n is the number of D numbers.

Step 6. verify the gained result by calculating CHEP: taking advantage of above process, the rank of alternative performers is obtained. In this paper, using Eqs. (15-16), the BPA of performers is obtained. Combining with the other two factors, "closeness in time" and "task relatedness", which are constant in order to control variables, to calculate the CHEP using Eqs. (1-7).

Take the above numerical data of performer A as an example. according to the Eqs. (15-16), the BPA of performer A is obtained:
$m_{performer}(5) = 0.7, m_{performer}(4) = 0.3$
Then assume
$m_{time}(5) = 0.2, m_{time}(5) = 0.6, m_{time}(5) = 0.2$
$m_{task}(5) = 0.2, m_{task}(4) = 0.5, m_{task}(3) = 0.3$
Thus according to the Eq. (5), the combination result $m_{performer} \bigoplus m_{time} \bigoplus m_{task}$ is:
$m_{performerA}(5) = m_{performer}(5) \bigoplus m_{time}(5) \bigoplus m_{task}(5) = 0.2373$
$m_{performerA}(4) = m_{performer}(4) \bigoplus m_{time}(4) \bigoplus m_{task}(4) = 0.7626$
According to the Eq. (7), we can conclude Betp(XD)=m(XD), because the used data is single subset.
where XD=CD,HD,MD,LD,and ZD.
Finally, the conditional human error probability (CHEP) $P(B|A)$ is calculated using Eq. (1) as: $P(B|A) = \sum_{XD} BetP(XD) \times P_{XD}(B|A) = 0.2373 \times \frac{1+0.01\times0}{1+0} + 0.7626 \times \frac{1+0.01\times1}{1+1} = 0.6224$ Assume that the failed probability of task B is 0.01. Given the previous failed task A, the reason why the failed probability of later task B is very high up to 0.6224 is that the relationship between the two successive tasks is complete dependent or highly dependent.

4 Application

In this section, a working model for post initiator HFEs of a nuclear power plant is used to illustrate the whole procedure of the proposed method. Furthermore, the working model of D numbers is used to determine the best performers that is a kind of multiple criteria decision making (MCDM). Then the traditional method in HRA is used to verify the proposed working model. At the end of this section, the effect of D numbers and its relationship with HRA will be discussed.

4.1 Selection of best performers

The identified input factors of interest and their relationship in the working model are shown in Fig. 3. According to the THERP method, three influential factors ("the average age difference", "the average trainees time difference" and "the average operation time in solving the same problem") directly affect the performers, which is the one of influential factor to dependence level. These factors are not necessarily relevant, but the model has enough complexity to illustrate the application of the methodology.

For each influential factor, anchor points and linguistic judgements corresponding to factor are provided by experts, as shown in Tables (1-4). In this paper, in order to figure out the best performer, D numbers is used as the representation of judgement offered by experts.

The representation of judgement in D numbers mainly depends on the number of experts who are in favor of specified judgment. Take the data in Table 5 as an example, as it clearly

shows, there are five experts giving judgement. And MD of one expert, CD of two experts, HD of two experts. Thus according to the percentage, it can obtain:
D(factor)=(HD,0.4)(CD,0.4)(MD,0.2)

Expert	Expert1	$Expert2$	$Expert3$	$Expert4$	$Expert5$
Level of factor	HD	HD	CD	CD	MD

Table 5: The example of judgement on factor

According to the above discussion, the representation of judgement in D numbers is shown in Table 6.

AlternativeD numbersFactor	Age	Operation	Time
$Performer\,1$	(5,0.4)(4,0.6)	(5,0.7)(4,0.3)	(5,1.0)
$Performer\,2$	(5,0.1)(4,0.6)(3,0.3)	(4,0.7)(3,0.3)	(5,0.1)(4,0.9)
$Performer\,3$	(4,0.3)(3,0.7)	(3,0.6)(2,0.4)	(4,0.3)(3,0.4)(2,0.3)
$Performer\,4$	(3,0.6)(2,0.4)	(3,0.1)(2,0.9)	(2,0.4)(1,0.6)

Table 6: The representation of influential factors in D numbers

The performers are determined by three factors (age, operation, time) jointly, thus this step is to aggregate the D numbers of the influential factors. Using Eqs. (8-13) , the result is shown in the Table 7:

Performer	$D \quad numbers$
Performer A	$(4.5, \frac{4.9}{16})(4.75, \frac{6}{16})(5, \frac{5.1}{16})$
Performer B	$(3.5, \frac{5.1}{30})(3.75, \frac{6.4}{30})(4, \frac{7.3}{30})(4.25, \frac{7.7}{30})(4.5, \frac{2.2}{30})(4.75, \frac{1.3}{30})$
Performer C	$(2.25, \frac{2.3}{24})(2.5, \frac{3.2}{24})(2.75, \frac{4.8}{24})(3, \frac{3.6}{24})(3.25, \frac{4.8}{24})(3.5, \frac{3.2}{24})(3.75, \frac{2.1}{24})$
Performer D	$(1.5, \frac{3.7}{20})(1.75, \frac{4.4}{20})(2, \frac{6}{20})(2.25, \frac{3.6}{20})(2.5, \frac{2.3}{20})$

Table 7: The integration of D numbers: $D_{age} \bigoplus D_{operation} \bigoplus D_{time}$

In order to select the optimal performer, given the existed judgement, the I(D) like a kind of utility needed to be calculated as a evaluate criteria by using Eq. (14).
$I(D_{performerA}) = 4.5 \times \frac{4.9}{16} + 4.75 \times \frac{6}{16} + 5 \times \frac{5.1}{16} = 4.735$
$I(D_{performerB}) = 3.5 \times \frac{5.1}{30} + 3.75 \times \frac{6.4}{30} + 4 \times \frac{7.3}{30} + 4.25 \times \frac{7.7}{30} + 4.5 \times \frac{2.2}{30} + 4.75 \times \frac{1.3}{30} = 3.995$
$I(D_{performerC}) = 2.25 \times \frac{2.3}{24} + 2.5 \times \frac{3.2}{24} + 2.75 \times \frac{4.8}{24} + 3 \times \frac{3.6}{24} + 3.25 \times \frac{4.8}{24} + 3.5 \times \frac{3.2}{24} + 3.75 \times \frac{2.1}{24} = 2.869$
$I(D_{performerD}) = 1.5 \times \frac{3.7}{20} + 1.75 \times \frac{4.4}{20} + 2 \times \frac{6}{20} + 2.25 \times \frac{3.6}{20} + 2.5 \times \frac{2.3}{20} = 1.955$

As shown in Fig. 4, performer A has biggest I(D), which means the worst performer. The assumption in this paper is that the best performer should be with the lowest CHEP. The relationship between I(D) and conditional human error probability (CHEP) is positive correlation, thus we conclude the performer D should have the lowest CHEP. In the next section, the verification will be discussed.

4.2 Verification

In the above section,the best performer has been selected. But for verifying our assumption, the final CHEP is necessary. By comparing the conditional human error probability (CHEP), the performer with the lowest CHEP means that performer has lowest probability to make mistakes.

In reality, the alternative performers will face same problem, under the same environment, at the same time, for which we assume the factors of "closeness in time" and "task relatedness" are same for each case. And the BPA of "similarity of performer" is obtained by Eqs. (15-16) and is illustrated in Table 8.

	Performers	Task	Time
A	m(5)=0.7;m(4)=0.3	m(5)=0.2;m(4)=0.6;m(3)=0.2	m(5)=0.2;m(4)=0.5;m(3)=0.3
B	$m(5)=\frac{0.2}{3}; m(4)=\frac{2.2}{3}; m(3)=\frac{0.6}{3}$	m(5)=0.2;m(4)=0.6;m(3)=0.2	m(5)=0.2;m(4)=0.5;m(3)=0.3
C	$m(4)=\frac{0.6}{3}; m(3)=\frac{1.7}{3}; m(2)=\frac{0.7}{3}$	m(5)=0.2;m(4)=0.6;m(3)=0.2	m(5)=0.2;m(4)=0.5;m(3)=0.3
D	$m(3)=\frac{0.7}{3}; m(2)=\frac{1.7}{3}; m(1)=\frac{0.6}{3}$	m(5)=0.2;m(4)=0.6;m(3)=0.2	m(5)=0.2;m(4)=0.5;m(3)=0.3

Table 8: The BPA of different cases

In this paper, it assumes that the weight of factors of same level is same. By using D-S evidence theory Eq. (6), the fused BPA of each factor is obtained as shown in Table 9.

Case	A	B	C	D
BBA	m(5)=0.2373 m(4)=0.7627	m(5)=0.0114 m(4)=0.9375 m(3)=0.0511	m(4)=0.6383 m(3)=0.3617	m(3)=1

Table 9: The final result of BPA

Due to the data used in the paper are all single subset as well as $m(\emptyset) = 0$, according to the Eq. (7), the numerical value of BetP is equal to the numerical value of BPA. Assume that basic human error probability of the subsequent task T_b is P(B)=0.01. Then the conditional human error probability (CHEP) is calculated using Eq. (1), the result is shown in Table 10 and Fig. 4(b).

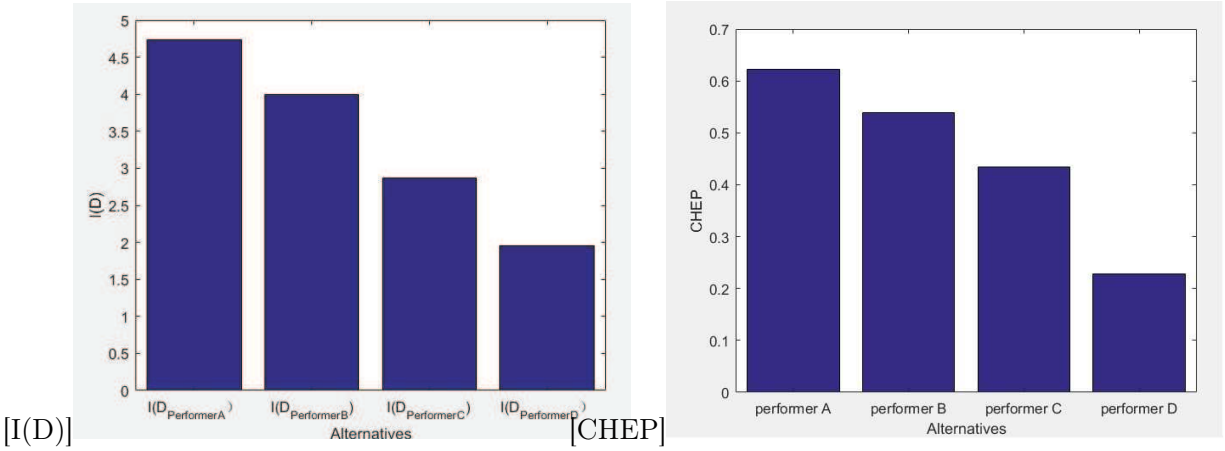
Case	A	B	C	D
CHEP	0.6224	0.5387	0.4337	0.2286

Table 10: The final result of CHEP

4.3 Discussion

The result of proposed method consists of two parts: the selection of alternatives and the calculation of CHEP. In the first part, the D numbers is used to represent the judgements given by experts. It is reasonable and intelligible way to handle such problem. The number of I(D) is treated as a criteria for performers. The smaller I(D) is, the better performer is. And the second part discusses the conditional human error probability (CHEP) that means that the failure probability of the following task given the failed preceding task. It is obvious the lower CHEP indicates the better situations. By comparing with Fig. 4, it concludes the relationship between I(D) and CHEP is positive correlation. Thus the gained result complies with our assumption, the performer A selected through proposed method has the highest CHEP 0.6224. It means that the task tends to fail with the most high probability when performer A is selected.

$I(D_{performerA}) > I(D_{performerB}) > I(D_{performerC}) > I(D_{performerD})$ $CHEP : A > B > C > D$.



[I(D)] [CHEP]

Figure 4: The comparison of ID and CHEP

5 Conclusion

In previous works in human reliability analysis (HRA), the influence of performer is not paid enough attention. In this paper, the evaluation method based on D numbers is presented to deal with the selection of alternative performers in an reasonable and simple way. In the real world, there is no doubt that performers are the main factor that affect our selection. The contribution of this paper is that offering a new way to translate the BPA into D numbers, which is more flexible than traditional ways and reduce the subjectivity at some extent. The final result demonstrates the effectiveness of the proposed method. In the future works, we would take the weights of different influential factors into consideration to make the predicted result more accurate.

Funding

The work is partially supported by National Natural Science Foundation of China (Grant Nos. 61573290, 61503237).

Conflict of interest

The authors declare that there is no conflict of interests regarding the publication of this paper.

Bibliography

[1] Carvalho, P. V.; Dos Santos, I. L.; Vidal, M. C. (2006). Safety implications of cultural and cognitive issues in nuclear power plant operation, *Applied Ergonomics*, 37(2), 211–223, 2006.

[2] Čepin, M. (2008). Importance of human contribution within the human reliability analysis (ijs-hra), *Journal of Loss Prevention in the Process Industries*, 21(3), 268–276, 2008.

[3] Chen, L.; Deng, X. (2018). A modified method for evaluating sustainable transport solutions based on ahp and Dempster–Shafer evidence theory. *Applied Sciences*, 8(4), Article ID 563, 2018.

[4] Chen, L.; Deng, Y. (2018). A new failure mode and effects analysis model using Dempster-Shafer evidence theory and grey relational projection method, *Engineering Applications of Artificial Intelligence*, 76, 13–20, 2018.

[5] Cooper, S. E.; Ramey-Smith, A.; Wreathall, J.; Parry, G. (1996). A technique for human error analysis (atheana). Technical report, Nuclear Regulatory Commission, Washington, DC (United States). Div. of Systems Technology; Brookhaven National Lab., Upton, NY (United States); Science Applications International Corp., Reston, VA (United States); NUS Corp., Gaithersburg, MD (United States).

[6] Cui, H.; Liu, Q.; Zhang, J.; Kang, B. (2019). An improved deng entropy and its application in pattern recognition, *IEEE Access*, 7, 18284–18292, 2019.

[7] Dempster, A. P. (2008). Upper and lower probabilities induced by a multivalued mapping, In *Classic Works of the Dempster-Shafer Theory of Belief Functions*, 57–72. Springer, 2008.

[8] Deng, X.; Jiang, W. (2019). D number theory based game-theoretic framework in adversarial decision making under a fuzzy environment, *International Journal of Approximate Reasoning*, 106, 194–213, 2019.

[9] Deng, X.; Jiang, W.; Wang, Z. (2019). Zero-sum polymatrix games with link uncertainty: A Dempster-Shafer theory solution, *Applied Mathematics and Computation*, 340, 101–112, 2019.

[10] Deng, Y. (2012). D numbers: Theory and applications, *Journal of Information & Computational Science*, 9(9), 2421–2428, 2012.

[11] Dong, Y.; Zhang, J.; Li, Z.; Hu, Y.; Deng, Y. (2019). Combination of evidential sensor reports with distance function and belief entropy in fault diagnosis, *International Journal of Computers Communications & Control*, 14(3), 293–307, 2019.

[12] Dzitac, I.; Filip, F. G.; Manolescu, M.-J. (2017). Fuzzy logic is not fuzzy: World-renowned computer scientist Lotfi A. Zadeh, *International Journal of Computers Communications & Control*, 12(6), 748–789, 2017.

[13] Fei, L.; Deng, Y. (2019). A new divergence measure for basic probability assignment and its applications in extremely uncertain environments, *International Journal of Intelligent Systems*, 34(4), 584–600, 2019.

[14] Fei, L.; Deng, Y.; Hu, Y. (2019). DS-VIKOR: A New Multi-criteria Decision-Making Method for Supplier Selection, *International Journal of Fuzzy Systems*, 21(1), 157–175, 2019.

[15] Gao, X.; Deng, Y. (2019). The generalization negation of probability distribution and its application in target recognition based on sensor fusion, *International Journal of Distributed Sensor Networks*, 15(5), 2019.

[16] Gao, X.; Deng, Y. (2019). The negation of basic probability assignment, *IEEE Access*, 7(1), 10.1109/ACCESS.2019.2901932, 2019.

[17] Gertman, D.; Blackman, H.; Marble, J.; Byers, J.; Smith, C. et al. (2005). The spar-h human reliability analysis method, *US Nuclear Regulatory Commission*, 2005.

[18] Guan, X.; Liu, H.; Yi, X.; Zhao, J. (2018). The Improved Combination Rule of D Numbers and Its Application in Radiation Source Identification, *Mathematical Problems in Engineering*, 2018.

[19] Han, Y.; Deng, Y. (2018). An evidential fractal ahp target recognition method, *Defence Science Journal*, 68(4), 367–373, 2018.

[20] Han, Y.; Deng, Y. (2019). A novel matrix game with payoffs of Maxitive Belief Structure, *International Journal of Intelligent Systems*, 34(4), 690–706, 2019.

[21] Jiang, W. (2018). A correlation coefficient for belief functions, *International Journal of Approximate Reasoning*, 103, 94–106, 2018.

[22] Jiang, W.; Huang, C. (2018). A multi-criteria decision-making model for evaluating suppliers in green SCM, *International Journal of Computers Communications & Control*, 13(3), 337–352, 2018.

[23] Jiang, W.; Wang, S. (2017). An uncertainty measure for interval-valued evidences, *International Journal of Computers, Communications & Control*, 12(5), 2017.

[24] Kang, B.; Deng, Y.; Hewage, K.; Sadiq, R. (2019). A method of measuring uncertainty for Z-number, *IEEE Transactions on Fuzzy Systems*, 27(4), 731–738, 2019.

[25] Kang, B.; Zhang, P.; Gao, Z.; Chhipi-Shrestha, G.; Hewage, K.; Sadiq, R. (2019). Environmental assessment under uncertainty using Dempster–Shafer theory and z-numbers, pages Published online, doi: 10.1007/s12652–019–01228–y, 2019.

[26] Kirwan, B. (1996). The validation of three human reliability quantification techniques – THERP, HEART and JHEDI: Part 1 – technique descriptions and validation issues, *Applied ergonomics*, 27(6), 359–373, 1996.

[27] Kontogiannis, T.; Malakis, S. (2009). A proactive approach to human error detection and identification in aviation and air traffic control, *Safety Science*, 47(5), 693–706, 2009.

[28] Lester, R. K.; McCabe, M. J. (1993). The effect of industrial structure on learning by doing in nuclear power plant operation, *The Rand Journal of Economics*, 418–438, 1993.

[29] Li, M.; Zhang, Q.; Deng, Y. (2018). Evidential identification of influential nodes in network of networks, *Chaos, Solitons & Fractals*, 117, 283–296, 2018.

[30] Li, X.; Chen, X. (2018). D-intuitionistic hesitant fuzzy sets and their application in multiple attribute decision making, *Cognitive Computation*, 10(3), 496–505, 2018.

[31] Li, Y.; Deng, Y. (2018). Generalized ordered propositions fusion based on belief entropy, *International Journal of Computers Communications & Control*, 13(5), 792–807, 2018.

[32] Li, Y.; Deng, Y. (2019). TDBF: Two Dimension Belief Function, *International Journal of Intelligent Systems*, 34, 10.1002/int.22135, 2019.

[33] Lin, S.; Li, C.; Xu, F.; Liu, D.; Liu, J. (2018). Risk identification and analysis for new energy power system in China based on D numbers and decision-making trial and evaluation laboratory (DEMATEL), *Journal of Cleaner Production*, 180, 81–96, 2018.

[34] Liu, H.; Dzitac, I.; Guo, S. (2018). Reduction of conditional factors in causal analysis, *International Journal of Computers, Communications & Control*, 13(3), 2018.

[35] Meng, D.; Liu, M.; Yang, S.; Zhang, H.; Ding, R. (2018). A fluid–structure analysis approach and its application in the uncertainty-based multidisciplinary design and optimization for blades, *Advances in Mechanical Engineering*, 10(6), https://doi.org/10.1177/1687814018783410, 2018.

[36] Pan, L.; Deng, Y. (2018). A New Belief Entropy to Measure Uncertainty of Basic Probability Assignments Based on Belief Function and Plausibility Function, *Entropy*, 20(11), 842, 2018.

[37] Paté-Cornell, E. (2012). On black swans and perfect storms: risk analysis and management when statistics are not enough, *Risk Analysis: An International Journal*, 32(11), 1823–1833, 2012.

[38] Pelusi, D.; Mascella, R.; Tallini, L. (2017). Revised gravitational search algorithms based on evolutionary-fuzzy systems, *Algorithms*, 10(2), 44, 2017.

[39] Pelusi, D.; Mascella, R.; Tallini, L. (2018). A fuzzy gravitational search algorithm to design optimal iir filters. *Energies*, 11(4), 736, 2018.

[40] Pelusi, D.; Mascella, R.; Tallini, L.; Nayak, J.; Naik, B.;Abraham, A. (2018). Neural network and fuzzy system for the tuning of gravitational search algorithm parameters, *Expert Systems with Applications*, 102, 234–244, 2018.

[41] Rong, H.; Ge, M.; Zhang, G.; Zhu, M. (2018). An approach for detecting fault lines in a small current grounding system using fuzzy reasoning spiking neural P systems, *International Journal of Computers Communications & Control*, 13(4), 521–536, 2018.

[42] Shafer, G. (1976). *A mathematical theory of evidence*, Vol. 42. Princeton University Press, 1976.

[43] Shankar, R.; Choudhary, D.; Jharkharia, S. (2018). An integrated risk assessment model: A case of sustainable freight transportation systems, *Transportation Research Part D: Transport and Environment*, 63, 662–676, 2018.

[44] Shorrock, S. T.; Kirwan, B. (2002). Development and application of a human error identification tool for air traffic control, *Applied ergonomics*, 33(4), 319–336, 2002.

[45] Song, Y.; Deng, Y. (2019). A new method to measure the divergence in evidential sensor data fusion, *International Journal of Distributed Sensor Networks*, 15(4), 2019.

[46] Su, X.; Li, L.; Shi, F.; Qian, H. (2018). Research on the fusion of dependent evidence based on mutual information, *IEEE Access*, 6, 71839–71845, 2018.

[47] Su, X.; Mahadevan, S.; Xu, P.; Deng, Y. (2015). Dependence assessment in human reliability analysis using evidence theory and AHP, *Risk Analysis*, 35(7), 1296–1316, 2015.

[48] Sun, R.; Deng, Y. (2019). A new method to identify incomplete frame of discernment in evidence theory, *IEEE Access*, 7(1), 15547–15555, 2019.

[49] Sun, R.; Deng, Y. (2019). A new method to determine generalized basic probability assignment in the open world, *IEEE Access*, 7(1), 52827–52835, 2019.

[50] Swain, A. (1964). Therp technique for human error rate prediction, In *Proceedings of the Symposium on Quantification of Human Performance, Albuquerque*, 1964.

[51] Swain, A. D. (1990). Human reliability analysis: Need, status, trends and limitations, *Reliability Engineering & System Safety*, 29(3), 301–313, 1990.

[52] Wang, J.; Qiao, K.; Zhang, Z. (2019). An improvement for combination rule in evidence theory, *Future Generation Computer Systems*, 91, 1–9, 2019.

[53] Xiao, F. (2018). A hybrid fuzzy soft sets decision making method in medical diagnosis, *IEEE Access*, 6, 25300–25312, 2018.

[54] Xiao, F. (2018). A novel multi-criteria decision making method for assessing health-care waste treatment technologies based on D numbers, *Engineering Applications of Artificial Intelligence*, 71(2018), 216–225, 2018.

[55] Xiao, F. (2019). Multi-sensor data fusion based on the belief divergence measure of evidences and the belief entropy, *Information Fusion*, 46, 23–32, 2019.

[56] Xiao, F. (2019). A multiple criteria decision-making method based on D numbers and belief entropy, *International Journal of Fuzzy Systems*, DOI: 10.1007/s40815–019–00620–2, 2019.

[57] Xiao, F.; Ding, W. (2019). Divergence measure of Pythagorean fuzzy sets and its application in medical diagnosis, *Applied Soft Computing*, 254–267, 2019.

[58] Xu, H.; Deng, Y. (2019). Dependent Evidence Combination Based on DEMATEL Method, *International Journal of Intelligent Systems*, 1555–1571, 2019.

[59] Yager, R. R. (2012). On z-valuations using zadeh's z-numbers, *International Journal of Intelligent Systems*, 27(3), 259–278, 2012.

[60] Yager, R. R. (2018). Fuzzy rule bases with generalized belief structure inputs, *Engineering Applications of Artificial Intelligence*, 72, 93–98, 2018.

[61] Yager, R. R.; Alajlan, N. (2017). Maxitive Belief Structures and Imprecise Possibility Distributions, *IEEE Transactions on Fuzzy Systems*, 25(4), 768–774, 2017.

[62] Yang, H.; Deng, Y.; Jones, J. (2018). Network division method based on cellular growth and physarum-inspired network adaptation. *International Journal of Unconventional Computing*, 13(6), 477–491, 2018.

[63] Yin, L.; Deng, Y. (2018). Toward uncertainty of weighted networks: An entropy-based model, *Physica A: Statistical Mechanics and its Applications*, 508, 176–186, 2018.

[64] Zadeh, L. A. (1965). Fuzzy sets. *Information and control*, 8(3), 338–353, 1965.

[65] Zadeh, L. A. (2019). Preliminary draft notes on a similarity-based analysis of time-series with applications to prediction, decision and diagnostics, *International Journal of Intelligent Systems*, 34(1), 107–113, 2019.

[66] Zheng, X.; Deng, Y. (2018). Dependence assessment in human reliability analysis based on evidence credibility decay model and iowa operator, *Annals of Nuclear Energy*, 112, 673–684, 2018.

[67] Zhou, M.; Liu, X.-B.; Chen, Y.-W.; Yang, J.-B. (2018). Evidential reasoning rule for MADM with both weights and reliabilities in group decision making, *Knowledge-Based Systems*, 143, 142–161, 2018.

[68] Zhou, M.; Liu, X.-B.; Yang, J.-B.; Chen, Y.-W.; Wu, J. (2019). Evidential reasoning approach with multiple kinds of attributes and entropy-based weight assignment, *Knowledge-Based Systems*, 163, 358–375, 2019.

[69] Zhou, M.; Liu, X.; Yang, J. (2017). Evidential reasoning approach for MADM based on incomplete interval value, *Journal of Intelligent & Fuzzy Systems*, 33(6), 3707–3721, 2017.

[70] Zhou, X.; Hu, Y.; Deng, Y.; Chan, F, T. S.; Ishizaka, A. (2018). A DEMATEL-Based Completion Method for Incomplete Pairwise Comparison Matrix in AHP, *Annals of Operations Research*, 271(2), 1045–1066, 2018.

[71] Zio, E. (2009). Reliability engineering: Old problems and new challenges, *Reliability Engineering & System Safety*, 94(2), 125–141, 2009.

[72] Zio, E.; Baraldi, P.; Librizzi, M.; Podofillini, L.; Dang, V. N. (2009). A fuzzy set-based approach for modeling dependence among human errors, *Fuzzy Sets and Systems*, 160(13), 1947–1964, 2009.

Author index