CONTENT

Vol. 6. No. 6-2020

Published bi-monthly since 2015.

ISSN 2664-066X (Print)
ISSN 2664-0678 (Online)

Publisher
Institute of Radio and
Information Systems (IRIS),
Vienna, Austria

Address:
1010 Wien, Austria,
Ebendorferstrasse 10/6b
media-publisher.eu/
synchroinfo-journal

APPROACH TO INTELLIGENT MONITORING OF CYBER ATTACKS

Alexey N. Nazarov
Expert ITU, Russia,
a.nazarov06@bk.ru

Dmitry V. Pantiukhin,
National Research University Higher School of Economics;
Moscow Institute of Physics and Technology, Moscow, Russia

Ilya M. Voronkov,
International Centre of Informatics and Electronics, ICIE;
National Research University Higher School of Economics, Moscow, Russia

Mikhail A. Nazarov,
CEO LLC "SmartTech", Moscow, Russia

DOI: 10.36724/2664-066X-2020-6-6-2-9

ABSTRACT

The results of many years of research on the subject of intellectual counteraction to cyberattacks are presented. Cloud solutions for the synthesis of the monitoring cluster of cyberattacks are based on the latest achievements with the use of neuron-fuzzy formalism. The main features of the synthesis of protection functions are determined and the features of the implementation of the security system of the object of risk in cyberspace are analyzed. Methodological approaches to solving the system problem of determining all ways of penetration of the attack on the object of risk and the formation of variants of their coatings are proposed. The peculiarities of applicability of the branch and boundary method for solving this problem are discussed.

KEYWORDS: *security function, cluster, method, Hadoop, neural network, monitoring.*

Information about authors

Alexey N. Nazarov, *Doctor of Technical Science, Professor, Expert ITU, Moscow, Russia*
Dmitry V. Pantiukhin, *Master of Science, Senior lecturer, National Research University Higher School of Economics; Lecturer, Moscow Institute of Physics and Technology, Moscow, Russia*
Ilya M. Voronkov, *Master of Science, Head of Departament, International Centre of Informatics and Electronics, ICIE; Invited lecturer, National Research University Higher School of Economics, Moscow, Russia*
Mikhail A. Nazarov, *CEO LLC "SmartTech", Moscow, Russia*

1. Introduction

For an arbitrary object of risk of an information and telecommunication system (ITCS) subjected to an information attack, in General, there is [1-3] a complete system (list) of security functions in the causal sense.

Table 1
Security Functions

Designation of security functions	Appointment of security functions
X_1	Preventing the occurrence of conditions conducive to the generation of (occurrence) destabilizing factors (hereinafter - DF)
X_2	Warning immediate manifestations of DF
X_3	Detection manifested DF
X_4	Prevention of exposure to risk in the manifested and revealed DF
X_5	Prevention of exposure to risk on the manifest, but the undetected DF
X_6	Detecting the impact of DF on the subject of risk
X_7	Localization (restriction) found the impact of DF on the subject of risk
X_8	Localization of undetected exposure to risk by DF
X_9	Dealing with the consequences of the localized impact of the detected object on the destabilizing factors risk
X_{10}	Dealing with the consequences of undetected localized exposure to risk by DF

The system of security functions allows to unite on the universal methodical basis and ways, means, technologies of information security from various, not overlapping on the physical, natural essence subject areas. Protection functions depend on a large number of destabilizing factors [1]. For specific risk objects, these functions are the basis of the security policy and are developed and investigated by the security service of the risk object. Attackers also analyze known security features of the object at risk to identify vulnerabilities. Each of the parties in relation to the object of risk pursues opposite goals. Protection functions are developed and implemented by each party on the basis of existing measures and means to ensure information security of risk objects. The essence of each of the protection functions listed in [1] is unique. And technologies of realization of functions of protection can and should be modified at reception of new knowledge about possibilities of information attack.

As shown in [1-3] in an information attack, the protection functions are interconnected by causal transitions.

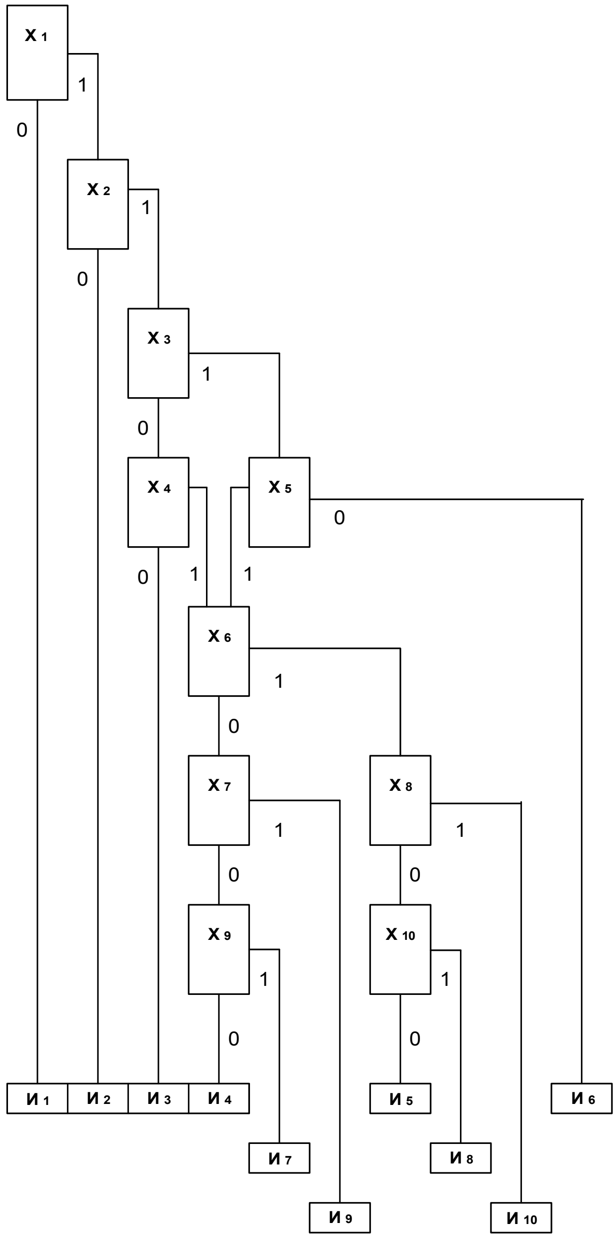

Fig. 1. The causal diagram of the security functions $X_1 \div X_{10}$ and results of attack $И_1 \div И_{10}$

Table 2
Final events in Fig. 1

$И_1 \div И_6$	Defence Provided
$И_7, И_8$	Defence Broken
$И_9, И_{10}$	Defence Destroyed

Therefore, the natural interpretation of such transitions is in the form of a digraph, which, in turn, can be represented as a matrix of incidents [7].

2. Aspects of security functions synthesis for any object of risk

General approaches to the practical creation of protection functions based on the fulfillment of the condition of reachability of an acceptable, sufficient level of security are given in [4]. In General, in order to protect the object, they develop an information security policy, guided by which they build barriers or boundaries of protection that prevent the possibility of carrying out an attack. As the boundaries of protection can be: protected area, access system to objects and buildings, user authentication, organization of password access to information of a certain category, etc. [3,7].

Each protection boundary can be represented as a directed graph [7], the set of vertices of which corresponds to the set of arguments (variables) of protection functions, and the set of arcs corresponds to the set of procedures for processing these variables. Such formalization naturally formalizes and concretizes the system of functions of protection of the object of risk.

Continuous automated monitoring of the risk object is possible on the basis of methods [4-6] that allow to quickly assess the level of security of the risk object, identify the attack and develop protection functions.

The task of further development of automated monitoring of the risk object is to simulate scenarios of trajectories of information attacks through the space in which the risk object of the infocommunication system is located. Which in turn is reduced to solving the problem of calculating the graph of the object of risk.

The solution of the problem fits into the following scheme of optimization stages:

– b sed on the graph model to overcome obstacles information attack to form all versions of attack success in relation to the risk object and the logical functions represented by the matrix of incidence (the achievability of the target of attack);

– crea e a variety of options for placing obstacles with the help of the problem of coverage;

– to carry out synthesis of variants of optimum placement by means of protection on the basis of the decision of a problem of conditional optimization;

– crea e additional coating options to improve the safety of certain critical elements;

– to carry out synthesis of additional variants of placement of means of protection providing different requirements of safety of elements of object of risk.

If the system of protection of the object of risk includes several consecutive boundaries, ordered in ascending order, then to access it it is necessary to overcome all the boundaries sequentially, starting with the lowest [7]. The security policy of the object of risk consists of a set of mechanisms for the protection of each obstacle.

Each protection mechanism is characterized by a vector of system characteristics, which include: the ability to quickly change the rules of interaction between users and data, the cost of its development and implementation, the required computing resources, etc.

At development or a choice of mechanisms of protection for some boundary of object of risk the means allocated for development and operation of this boundary of protection, losses in ITCS caused by success of attack, structure and qualification of developers, resources, structure and characteristics of a boundary etc.Thus, any mechanism of protection should be based on scientifically-practically proved methodical recommendations should be considered and methodically accurately applied.

The amount of options entry risk object specifies the number of access of the offender to the facility. And each penetration is prevented by interrupting the path at least in one edge of the graph. It is necessary to exclude (interrupt) all ways of penetration. This is the problem of finding the minimum cross-section on a graph, which is solved by determining the minimum coverage on the incidence matrix.

The formulation of the problem of covering - all paths of penetration to cover (interrupt) the minimum number of edges, in General, is reduced to the problem of conditional optimization. Thus it is possible to achieve that redundancy of not realized possibilities of covering edges of the graph tends to a minimum, provided that each edge covers at least one path.

This problem is solved by the method of branches and boundaries. From the point of view of system analysis, the process of obtaining all penetration paths and forming many variants of their coatings is the task of decomposing a complex problem into simple subtasks. After this task, according to the theory of system analysis, the problem of synthesis of optimal placement of hardware and software protection of ITCS is solved.

3. Monitoring clust

The authors analyzed the principles, approaches and technological procedures for the organization of monitoring from fairly General premises. Methodological approaches to the creation of algorithms and software solutions in the Hadoop web-programming environment in relation to a wide class of object monitoring tasks in the web-space are developed. For the first time, the topology of the hadoop monitoring cluster, which has a common application, is schematically shown in Fig. 3. Researches are conducted and algorithms of measurements of attributes of objects of monitoring in web-space taking into account requirements of unity of measurements are offered. System requirements for Hadoop monitoring cluster design have been developed.

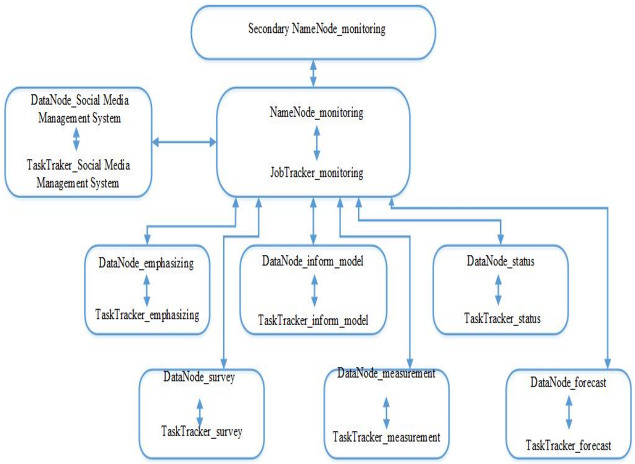

Fig. 3. Topology of the monitoring cluster Hadoop. Description daemons are given in [5]

The following hierarchical structure of the intelligent system of the ***TaskTraker*_**status daemon is proposed as a complex of software modules of the automated system (AS) for assessing the state of the monitoring object in the web space. By analogy with [5] daemon contains the following functional modules: system, fuzzy production rules, describing the work ID taking into account expert assessments; neuro-fuzzy network whose structure reflects the system of fuzzy production rules; clear self-learning neural network (NS) to solve problems of classification and clustering of input vectors. As noted in [3,4], this hierarchy has a common application and is therefore suitable for different monitoring objects based on the Hadoop monitoring cluster shown in Fig. 3.

The level of identification of changes in attributes (characteristics, parameters) of the monitoring object, intended for classification by the vector of signs of changes in the elements and nodes of the monitoring object formed by the sensors of these changes, is illustrated in Fig. 4.

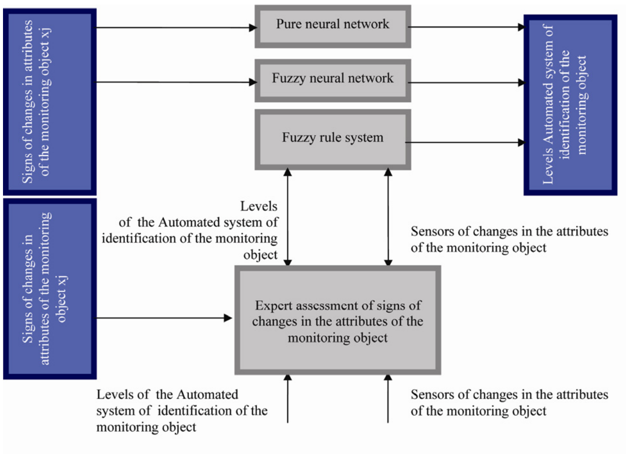

Fig. 4. Diagram of the adaptive classifier the level of identification of changes in the attributes of the monitoring object

The main system requirements for ***TaskTraker*_**status daemon of AS monitoring object in the web-space, the presence of which is mandatory:

– presentation of a priori experience of experts in web-monitoring of the selected object in the form of a knowledge base described by the system of production rules;

– availability of criteria base for decision-making on changing attributes of the monitoring object;

– fuzzy logical conclus on, allowing to use the experience of experts in web-monitoring of the selected object in the form of a system of fuzzy production rules for the initial configuration of the information field (system of interneuron connections) fuzzy NS;

– plug-in aggregation services and services for processing unstructured information about changes in the attributes of the monitoring object for further analysis;

– the ability of the NS to classification and clustering;

– the ability of the NS to extract knowledge about the profile and mechanism of implementation of changes in the attributes of the monitoring object in the web space;

– the ability of the information field of the national Assembly to gain experience in the process of learning and self-learning.

Software that meets the above requirements must be developed in the Hadoop environment. In addition, the daemon ***TaskTraker*_**status monitoring object in the web-space should be based on service-oriented integration methods in terms of scalability of its functional features.

4. The mechanism of fuzzy inference

This mechanism is based on the presentation of the experience of specialists (experts) on web-monitoring by the system of fuzzy production rules of the IF-THEN-type. For example:

Π_1: IF $\tilde{x}_1$ IS A_{11} AND … $\tilde{x}_n$ IS A_{1n}, THEN $\tilde{y}$ IS B_1;

Π_2: IF $\tilde{x}_1$ IS A_{21} AND … $\tilde{x}_n$ IS A_{2n}, THEN $\tilde{y}$ IS B_2;

…

Π_k: IF $\tilde{x}_1$ IS A_{k1} AND … $\tilde{x}_n$ IS A_{kn}, THEN $\tilde{y}$ IS B_k,

where $\tilde{x}_i$ and $\tilde{y}$ are fuzzy input and output variables respectively; A_{ij} and B_i, $j = 1,\ldots,n$, $i = 1,\ldots,k$, the corresponding membership functions.

Combining the capabilities of NS and fuzzy inference is one of the promising approaches to the organization of artificial intelligence systems. As it was shown in [5], fuzzy logic systems compensate for the main "opacity" of the NS: in the representation of knowledge and the ability to explain the results of the intellectual system, i.e. complement the NS. Fuzzy inference formalism works in the absence of knowledge about the attributes of monitoring objects and their changes for any monitoring objects, which is important when new attributes with unknown dynamics of changes appear.

For the functional demon ***TaskTraker*_**status monitoring in the web-space is a very important feature of neuro-fuzzy networks as the ability to automatically generate a system of fuzzy production rules in the process of teaching and learning, extracting hidden patterns from data, the input training samples.

Neural network training algorithms using stochastic properties of the dynamics of changes in the attributes of the monitoring object in the web space should be based on the standard method of minimizing the generalization error [3], on the basis of minimizing the quadratic residual functional on the training sample, searching for the gradient extremum of the target function of the error distribution density using the Robbins-Monroe procedure.

5. Hierarch levels and technological features of the functioning of the intelligent system daemon ***TaskTraker*_status**

The ability of NS to classify and cluster is used in the daemon to solve two main tasks:

1) classification of input ve or, for example, vector of signs of changes of attributes of object of monitoring;

2) xpansion of classification at appearance on an input of the classifier of the combination of signs of changes of attributes which was not earlier met.

Let the currently complete parcel space be $X = \{\widetilde{x}_1, \ldots, \widetilde{x}_m\}$ and a full space of conclusions $Y = \{\widetilde{y}_1, \ldots, \widetilde{y}_n\}$. Fuzzy causal relationships $\widetilde{x}_i \to \widetilde{y}_j$, $i = 1, \ldots, m$, $j = 1, \ldots, n$ between the elements of these spaces can be represented as a matrix R with elements r_{ij}, $i = 1, \ldots, m$, $j = 1, \ldots, n$, and the premises and conclusions between them can be presented in the form of: $B = A \bullet R$, where $\bullet$ – composition operation, for example, max-min-composition.

In fuzzy inference, expert knowledge $A \to B$ reflects a fuzzy relation $R = A \to B$, where the operation $\to$ corresponds to a fuzzy implication. A fuzzy relation R can be considered as a fuzzy subset of the direct product $X \times Y$ of a complete set X and inferences Y, and the process of obtaining a fuzzy inference B result by premise A and knowledge $A \to B$ can be considered as a compositional rule: $B = A \bullet R = A \bullet (A \to B)$.

From practice [5] it follows that at the level of experience accumulation, the neuro-fuzzy classifier of vectors of characteristics, parameters of change (dynamics of attributes) of the monitoring object should be designed in the form of a three-layer fuzzy NS (Fig. 5) with the ability to reduce (compress) the number of features.

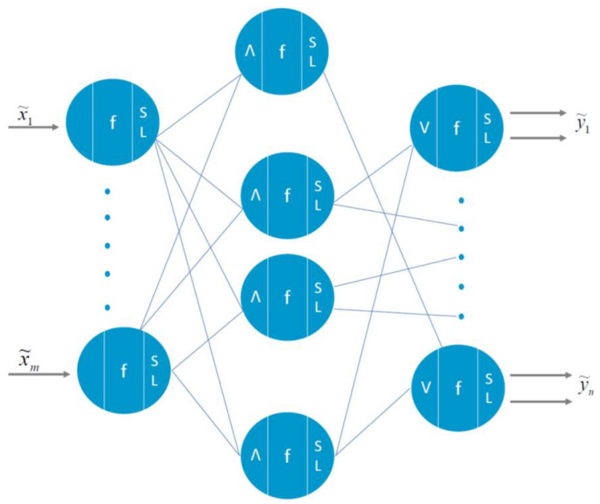

Fig. 5. Scheme of neuro-fuzzy classifier as S L-complementary pair of accessory functions

Each input vector from space can be matched with a fuzzy formal neuron (FN). The middle layer contains fuzzy FN performing a logical inference operation (for example, min) on combinations of fuzzy statements (FS) of the first NS layer to form a system of fuzzy classification conclusions.

The third layer of fuzzy NS is formed from fuzzy FN "OR" (by the number of fuzzy conclusions ,) and forms a vector of output fuzzy conclusions in accordance with the system of fuzzy rules set by experts.

5. xperimental results for application of convolution neural network for intrusions classification based on UNSW-NB15 dataset

For experimental estimation we use well-known UNSW-NB15 intrusion dataset [8]. This dataset contains both records about network connections, and the traffic itself (about 100GB) from the test computer network (3 servers). Dataset contains 9 types of attacks: Fuzzers, Analysis, Backdoors, DoS, Exploits, Generic, Reconnaissance, Shellcode, Worms. Each records contains 47 fields – nformation about network connections and 2 fields - information about the type of intrusion (or normal packets). The total number of records is about 2 million. Information about network connections is divided into 4 text files (csv format). Files for training and testing of classifiers, containing respectively 175341 and 82332 records, are separately presented

We introduce a short review of UNSW-NB15 dataset usage (taken from [9]). Firstly, in 2015, creators of UNSW-NB15 dataset propose and compare significant feature selection approach to UNSW-NB15 and KDD99 datasets and build classifiers based on Naive Bayes and EM-clustering approaches [8]. They show that such simple approaches works with accuracy about 30-40% but some classes are not recognized at all. In [11] (2017) other authors combine Random forest classification method and feature selection approach, and reach better results both for KDD99 and UNSW-NB15 dataset. In same 2017, another author [12] combine Random Forest method with “Logitboost” [13] boosting algorithm and show better results on both datasets compared to clear Random forest methods [14] and some previous approaches. Boosting is quite attractive approach to create classifiers, as example in [16] authors compare a set of boosting methods (Bagged Tree, AdaBoost, GentleBoost, LogitBoost and RUSBoost algorithms) on UNSW-NB15 dataset and show that Bagged Tree and GentleBoost classifiers show superior performance (see [16] for details). Another comparison of boosting presented in [17] and ensembling methods in [18].

In 2016 researches apply genetic search to select appropriate features to each class and use Support Vector Machine to classify intrusions with such features [19]. They achieve high accuracy, more than 90% for all classes except of “Exploits” (~80%), and false positive rates less then 0.1%.

In 2018 quite interesting approach was introduced in [20]. Authors use multiscale wavelet transform and perceptron-like neural network with Hebbian learning for anomaly detection in records with HTTP protocol. They achieve Mean Accuracy 93.56% (wherein Mean TPR is 73.55%, Mean FPR is 4.46%, Mean TNR is 95.53%, Mean FNR: is 26.44%).

Creators of UNSW-NB15 continues their work and propose a Collaborative Anomaly Detection Framework, which is based on modification of Gaussian Mixture Model [21]. This system installed on each node of cloud network and each node in training phase create a statistical model of normal data by approximation of probability distribution function using Gaussian Mixture Model. Then, in testing phase, such model can be used to detect an anomaly, which is interpreted as intrusion.

Such approach can be applied only to detection but not to classification of intrusion. Authors achieve results 96% in accuracy and detection rate and 4-9% in false positive rate. Similar approaches but with Hidden Markov Models also was applied in [22], with Beta Mixture Model in [23], with Dirichlet Mixture Mechanism in [24].

Some comparison of machine learning approaches for intrusion classification over UNSW-NB15 dataset are made in [25]. Authors compare Support Vector Machine, Multilayer Perceptron, Restricted Boltzmann Machine, Sparse Autoencoder and deep learning architecture with embedding (like word2vec approach).

They show that deep learning approach is better in average than others, but, unfortunately, does not provide information about performance on each classes for UNSW-NB15 dataset. Autoencoders is also interesting approach to intrusion detection, in [26] authors present a two stage approach with autoencoders then second stage uses a results (score – output of classification unit) from first stage. Such approach shows follow results: 89.134% in accuracy and a 0.7495% in FAR for the UNSW-NB15 dataset. Combination of deep autoencoder, Support Vector Machine and Artificial Bee Colony searching method was used in [27] and show detection accuracy about 90% and FAR about 5%.

In [28] authors provide using a Long-Short-Term-Memory (LSTM) neural network, which is a type of Recurrent Neural Network, for intrusion detection over UNSW-NB15 dataset. They achieve quite high results (Precision=98.02%; Accuracy=99.41%; TPR=97.97%; TNR=99.53%, FNR=2.03%, and FPR=0.47%) in detection and show that such approach works better than, for example, Support Vector Machine. Bidirectional LSTM was used in [29] and show average 85% in precision and 88% in recall metrics. It is also shown that some classes are not recognized at all, due to imbalance amount of data.

In [30] deep learning architecture (16 layers, including fully connected with ReLU activation, dropout, and batch normalization layers) was created and tested on various intrusion datasets, including UNSW-NB15. Results not so high as in previous researches, but this architecture in used for large set of different datasets, and clearly show a problem of imbalanced classes.

Authors of [31] provide a set of techniques (Bootstrap Aggregation, Synthetic Minority Over-sampling, Undersampling, and Class Balancer) to deal with imbalanced classes in intrusion detection system. They show minor advantages in term of area under ROC-curve to whole dataset, but, unfortunately, did not provide any information for classes.

Latest article [32] use a multiple-layer approach consisting of a coarse layer and a fine layer, in which the coarse layer with the deep convolutional neural network model focuses on identification of N abnormal classes and a normal class. While in the fine layer, an improved model based on gcForest (caXGBoost) further classifies the abnormal classes into N-1 subclasses. The proposed framework has been compared with the existing deep learning models using dataset NBC, a combination of UNSW-NB15 and CICIDS2017 including 101 classes. The experimental results show that method outperforms other single deep learning methods in terms of accuracy, detection rate and FAR and works well with imbalanced classes.

We study that for intrusion detection tasks it is a common situation then amount of examples of data of various classes is quite different (e.g. Normal class has 2 millions records, Worms class – only 174).

In this case training process of neural networks for intrusion recognition need to be modified to achieve better recognition of classes with small amount of examples. We use an approach of class-balanced batch forming and show in experiment that it can improve recognition performance of classes with small number of examples by the expense of decreased recognition performance of classes with large number of examples [9].

We use a convolution neural network for build a classifier. It consist from 5 convolutional layers with 'sigmoid' activations and 3 fully connected layers, first and second with 'sigmoid' and last with 'softmax' activation. Number of inputs – 190, number of outputs – 10 (9 intrusions classes and Normal). UNSW-NB15 has 47 features field, some of them relates only to that computer network which used for dataset collection, so we drop 7 irrelevant features. We code categorical features using one-hot encoding scheme. Therefore, for categorical features we have output vector of 153 (129+8+16) length. 15 of 37 numerical features scaled by division to appropriate value to make range almost equal and remaining 22 features unchanged.

Neural network trained using 'RMS-prop' [10] method, which is one of the modification of gradient-based method realized in 'keras' [10] library. Training is an iterative process in which loss-function minimizes. On each iteration, examples of inputs and corresponding desired outputs processed in neural network. Common choice for large datasets is to use so called 'batches' – combination of inputs\outputs of some length. This batches varies in each iteration and random input\output examples taken to a batch. In our work, we show that for situations then amount of examples in classes is very different better to use another approach for batch forming, namely - use predefine amount of examples of each class in batch. For example, if batch size is 300 and we have 10 classes then we can form a batch that consist of 30 random examples of each class.

As a result, we show that applying a class-balanced approach it is possible to drastically increase a recognition rate for classes with small number of examples corresponding to imbalanced training [9]:

– from 0% to 38% for Worms class (174 examples),

– rom 44% to 79% for Shellcode class (1511 examples),

– f om 0% to 53% for Backdoor class (2329 examples),

– f om 10% to 13% for Analysis class (2677 examples),

– f om 73% to 82% for Reconnaissance class (13987 examples),

– from 13% to 56% for DoS class (16353 examples),

but decrease recognition rate for classes with large number of examples:

– fro 77% to 53% for Fuzzers class (24246 examples),

– rom 81% to 56% for Exploits class (44525 examples),

– rom 98% to 97% for Generic class (215481 examples),

– fro 97% to 71% for Normal class (2218761 examples).

So, such approach can be applicable only for classes with small number of examples. Approach can be enlarged to control relative importance of classes by varying proportion of classes presented in batch.

Conclusion

Theoretical and technological research, which have a common application, and allow the machine to develop intelligent cloud solutions for the synthesis of monitoring systems and protection against cyber attacks on the basis of the protection functions of the object of risk on the basis of the following new results.

A machine-oriented graph approach to the synthesis of protection functions is developed, based on the graph representation of attack scenarios and a new application of the branch and boundary method. As a result of the application of this method, graph paths to the success of a cyber attack are suppressed.

For the topology of the monitoring a Hadoop cluster developed and tested guidelines for the synthesis of the demon Taktakishvili responsible for the task of assessing the status of the object of observation and identification information model, with the features of cloud computing. The principles and approaches based on neuro-fuzzy solutions that can be used as a basis for the design of intelligent automated systems for monitoring objects in the web space are proposed.

The mechanisms of decision-making based on the formalization of a priori experience of experts in the fuzzy base of fuzzy production rules are proposed. In the framework of solving the problems of classification and extension of classification of input data on the characteristics of the dynamics of attributes of the object of monitoring, the possibilities of a neuro-fuzzy classifier in the form of a three-layer fuzzy neural network are investigated.

Acknowledgement

Support from the Basic Research Program of the National Research University Higher School of Economics is gratefully acknowledged.

References

1. A. N. Nazarov. “Estimation of information safety level of modern infocommunication networks on basis of logic-probability approach,” in *Journal of Automation and Remote Control*, Vol. 68. Issue 7, 2007, pp. 1165-1176.

2. A. N. Nazarov. “Logical-and-probabilistic model for estimating the level of information security of modern information and communication networks,” *Journal of Telecommunications and Radio Engineering*, USA, 2010. Vol. 69. No. 16, pp. 1453-1463.

3. A. Nazarov & K. Sychev 2011, Models and methods for calculating the indicators of quality of functioning of the equipment units and structural parameters of the network the next generation networks, 2th edn, LLC Policom, Russia, Krasnoyarsk, 491 p.

4. A. Nazarov 2016, “Assessment of security from information attacks”, *Telecommunications*, no. 5, pp. 23-33.

5. A. Nazarov, M. Nazarov, D. Pantiuhin, S. Pokrova & A Sychev 2015, “Automation of monitoring processes in web-based neuro-fuzzy formalism”, T-Comm. Vol. 9. No. 8, pp. 26-33.

6. A Nazarov 2017, “Syntez of security functions against cyber-attacks”, T-Comm. Vol. 11. No. 9, pp. 80-85.

7. V Kostin 2017, “Synthesis of the optimal placement of hardware for physical protection systems for critical facilities”, Information Technology. Vol. 23. No. 1, pp. 41-49.

8. N. Moustafa, J. Slay, 2016 “The evaluation of network anomaly detection systems: statistical analysis of the UNSW-NB15 data set and the comparison with the KDD99 data set,” *Information Security Journal: A Global Perspective*. Vol. 25. No. 1-3, pp. 18-31.

9. D. Pantiukhin, A. Nazarov, I. M. Voronkov, 2019. “Intelligent methods for intrusion detection in local area networks”, *Proceedings of the 6th International Conference on Actual Problems of System and Software Engineering (APSSE 2019)* / Ed. by B. Pozin, A. R. Cavalli, A. Petrenko. P. 138-149.

10. “Keras: The Python Deep Learning Library.” [Online]. Available: https://keras.io.

11. T. Janarthanan and S. Zargari, "Feature selection in UNSW-NB15 and KDDCUP'99 datasets," 2017 IEEE 26th International Symposium on Industrial Electronics (ISIE), Edinburgh, 2017, pp. 1881-1886. doi: 10.1109/ISIE.2017.8001537

12. M. H. Kamarudin, C. Maple, T. Watson and N. S. Safa, "A LogitBoost-Based Algorithm for Detecting Known and Unknown Web Attacks," *IEEE Access*, vol. 5, pp. 26190-26200, 2017. doi: 10.1109/ACCESS.2017.2766844

13. J. Friedman, T. Hastie, and R. Tibshirani, “Additive logistic regression: A statistical view of boosting,” Ann. Stat., vol. 28, no. 2, pp. 337-374, 2000.

14. H. Tribak, B. L. Delgado-Márquez, P. Rojas, O. Valenzuela, H. Pomares, and I. Rojas, “Statistical analysis of different artificial intelligent techniques applied to intrusion detection system,” *Proc. Int. Conf. Multimed. Comput. Syst.*, 2012, pp. 434-440.

15. R. Primartha and B. A. Tama, "Anomaly detection using random forest: A performance revisited," *2017 International Conference on Data and Software Engineering (ICoDSE)*, Palembang, 2017, pp. 1-6. doi: 10.1109/ICODSE.2017.8285847

16. V. Timčenko and S. Gajin, "Ensemble classifiers for supervised anomaly based network intrusion detection," *2017 13th IEEE International Conference on Intelligent Computer Communication and Processing (ICCP),* Cluj-Napoca, 2017, pp. 13-19. doi: 10.1109/ICCP.2017.8116977

17. B. Patel, Z. Somani, S. A. Ajila and C. Lung, "Hybrid Relabeled Model for Network Intrusion Detection," *2018 IEEE International Conference on Internet of Things (iThings) and IEEE Green Computing and Communications (GreenCom) and IEEE Cyber, Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData)*, Halifax, NS, Canada, 2018, pp. 872-877. doi: 10.1109/Cybermatics_2018.2018.00167

18. N. Moustafa, B. Turnbull and K. R. Choo, "An Ensemble Intrusion Detection Technique Based on Proposed Statistical Flow Features for Protecting Network Traffic of Internet of Things," *IEEE Internet of Things Journal*. Vol. 6. No. 3, pp. 4815-4830, June 2019. doi: 10.1109/JIOT.2018.2871719

19. H. Gharaee and H. Hosseinvand, "A new feature selection IDS based on genetic algorithm and SVM," *2016 8th International Symposium on Telecommunications (IST),* Tehran, 2016, pp. 139-144. doi: 10.1109/ISTEL.2016.7881798

20. S. Siddiqui, M. S. Khan and K. Ferens, "Multiscale Hebbian neural network for cyber threat detection," *2017 International Joint Conference on Neural Networks (IJCNN),* Anchorage, AK, 2017, pp. 1427-1434. doi: 10.1109/IJCNN.2017.7966020

21. N. Moustafa, G. Creech, E. Sitnikova and M. Keshk, "Collaborative anomaly detection framework for handling big data of cloud computing," *2017 Military Communications and Information Systems Conference (MilCIS),* Canberra, ACT, 2017, pp. 1-6. doi: 10.1109/MilCIS.2017.8190421

22. N. Moustafa, E. Adi, B. Turnbull and J. Hu, "A New Threat Intelligence Scheme for Safeguarding Industry 4.0 Systems," *IEEE Access*, vol. 6, pp. 32910-32924, 2018. doi: 10.1109/ACCESS.2018.2844794

23. N. Moustafa, J. Slay and G. Creech "Novel Geometric Area Analysis Technique for Anomaly Detection using Trapezoidal Area Estimation on Large-Scale Networks," *IEEE Transactions on Big Data*. doi: 10.1109/TBDATA.2017.2715166

24. N. Moustafa, K. R. Choo, I. Radwan and S. Camtepe, "Outlier Dirichlet Mixture Mechanism: Adversarial Statistical Learning for Anomaly Detection in the Fog," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 8, pp. 1975-1987, Aug. 2019. doi: 10.1109/TIFS.2018.2890808

25. J. Yan, D. Jin, C. W. Lee and P. Liu, "A Comparative Study of Off-Line Deep Learning Based Network Intrusion Detection," *2018 Tenth International Conference on Ubiquitous and Future Networks (ICUFN)*, Prague, 2018, pp. 299-304. doi: 10.1109/ICUFN.2018.8436774

26. F. A. Khan, A. Gumaei, A. Derhab and A. Hussain, "A Novel Two-Stage Deep Learning Model for Efficient Network Intrusion Detection," *IEEE Access*, vol. 7, pp. 30373-30385, 2019. doi: 10.1109/ACCESS.2019.2899721

27. Q. Tian, J. Li and H. Liu, "A Method for Guaranteeing Wireless Communication Based on a Combination of Deep and Shallow Learning," *IEEE Access*, vol. 7, pp. 38688-38695, 2019. doi: 10.1109/ACCESS.2019.2905754

28. S. Xiao, J. An and W. Fan, "Constructing an Intrusion Detection Model based on Long Short-term Neural Networks," *2018 IEEE/ACIS 17th International Conference on Computer and Information Science (ICIS)*, Singapore, 2018, pp. 355-360. doi: 10.1109/ICIS.2018.8466445

29. S. Yang, "Research o etwork Behavior Anomaly Analysis Based on Bidirectional LSTM," *2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, Chengdu, China, 2019, pp. 798-802. doi: 10.1109/ITNEC.2019.8729475

30. R. Vinayakumar, M. Alazab, K. P. Soman, P. Poornachandran, A. Al-Nemrat and S. Venkatraman, "Deep Learning Approach for Intelligent Intrusion Detection System," *IEEE Access*, vol. 7, pp. 41525-41550, 2019. doi: 10.1109/ACCESS.2019.2895334

31. C. Wheelus, E. Bou-Harb and X. Zhu, "Tackling Class Imbalance in Cyber Security Datasets," *2018 IEEE International Conference on Information Reuse and Integration (IRI),* Salt Lake City, UT, 2018, pp. 229-232. doi: 10.1109/IRI.2018.00041

32. X. Zhang, J. Chen, Y. Zhou, L. Han and J. Lin, "A Multiple-layer Representation Learning Model for Network-Based Attack Detection," *IEEE Access*. doi: 10.1109/ACCESS.2019.2927465

THREE-DIMENSIONAL PIEZOELECTRIC ACCELEROMETER FOR MEASURING DYNAMIC PARAMETERS OF MOVING OBJECTS

Tofig M. Mansurov, *Azerbaijan Technical University, Baku, Azerbaijan*
Chingiz P. Aliyev, *Ministry of Defense Industry, Baku, Azerbaijan*

DOI: 10.36724/2664-066X-2020-6-6-10-16

ABSTRACT

Analysis of structural construction of existing three-dimensional accelerometers is carried out, in which due to the opposite direction of sensitivity vectors and direction of measured inertial force, they allow to measure only linear inertia and accelerations varying in relatively small limits. On the other hand, due to the presence of elements in them that create electromagnetic fields, their construction is somewhat complicated. In addition, in known accel-erometers, due to the absence of a measurement object position sensor and an electronic control circuit, the func-tionality of the accelerometer is limited, Due to lack of integrator for acceleration integration, speed measurement is not provided, Due to rigid attachment of piezoelectric sensing elements through their bases, during measurement they generate interference signals, The amplitude of which exceeds the amplitude of the useful signal, which does not ensure reliability of the measured acceleration or speed and due to the presence of sensitive piezoelectric ele-ments, Having higher rigidity and requiring relatively large mechanical forces for generation of signals during mo-tion of acceleration and speed measurement object in automatic mode, which reduces sensitivity of accelerometer.

Invention proposes new design and control scheme of three-dimensional piezoelectric accelerometer for measure-ment of dynamic parameters of moving objects in automatic mode. Wherein providing the position sensor of the moving measurement object with a three-axis signal detection unit to determine a direction of motion with meas-urement of acceleration or linear velocity when the measurement object moves along the coordinate axis; With the help of differential operational amplifiers through pulse generators and integrators, Speed measurement is provided by damping piezoelectric elements when the measurement object moves along one of the coordinate axes; Across the other two axes, the generated interference signals by amplitude are significantly reduced by their redemption; Having a piezoelectric element in the structure in the form of two-layer flat plates with an excitation section and a sensor section; Which generates signals at fast-changing acceleration and speed of moving object, amplitude of signal and sensitivity increases in 4-5 due to generation of signal by sensitive elements in vibration excitation mode created in accelerometer.

Mathematical basis for accelerometer control circuits, mathematical model of its functioning is developed, applica-tion of which increases sensitivity of measurement of linear accelerations and speeds of moving objects and expands functional capabilities.

KEYWORDS: *accelerometer, piezoelectric element, sensitivity vector, inertial force, linear acceleration, velocity, vibration measurement*

Information about authors

Tofig M. Mansurov, *Professor of department "Telecommunications and Information Security", Doctor of Technical Sciences, Azerbaijan Technical University, Baku, Azerbaijan*

Chingiz P. Aliyev, *Director General of Production Association "UFUG", Ministry of Defense Industry, Baku, Azerbaijan*

Introduction

The relevance of oscillatory processes occurring study in moving objects is associated with the development of a three-dimensional accelerometer for measuring dynamic parameters of moving objects in real time. The use of information-measuring systems makes it possible to automate process of measuring the dynamic parameters of moving objects, which facilitates accounting of rare events that strongly affect the measurement results [2, 3, 6, 7].

For this purpose, three-coordinate position sensors are installed in the measuring devices, which perceive movements of measurement object along three coordinates of a rectangular coordinate system, placed on moving objects and converting a mechanical quantity into a corresponding electrical signal. This signal goes to the recording device. Piezoelectric vibration measuring transducers, which are generator-type sensors, which convert vibration acceleration into a proportional electrical signal, are mainly used as such position sensors.

The accelerometer consists of a sensitive mass fixed in an elastic suspension, under the influence of the apparent acceleration, a deviation from the initial position occurs, proportional to this acceleration [1-3]. Accelerometers are devices with one degree of freedom and contain an inertial mass, flat springs as an elastic support system, and a damping device. Position sensors detect the movement of the inertial mass relative to the measurement object and convert it into a proportional electrical signal. In this regard, another component of the accelerometer is a displacement detector capable of measuring the amplitudes of vibration vibrations in a certain range or linear acceleration. The output signal of the accelerometer, which is proportional to the value of the acceleration of mechanical vibrations, is converted into an electrical signal by means of a meter.

The object of research is a three-dimensional piezoelectric accelerometer for measuring the dynamic parameters of moving objects in an automatic mode.

The subject of research is study of basic principles and patterns of increasing sensitivity and reliability and expanding the functionality of a three-dimensional piezoelectric accelerometer for measuring dynamic parameters of moving objects in an automatic mode.

The aim of the work is to improve the accuracy and reliability of measurements, sensitivity and expand the functionality of a three-dimensional piezoelectric accelerometer for measuring dynamic parameters of moving objects with a rapidly changing direction of movement of measurement object along the three axes "XYZ" of a rectangular coordinate system in automatic mode.

Formulation of the problem

To measure the dynamic parameters of moving objects in automatic mode, piezoelectric accelerometers are used as a working element of the accelerometer. The use of piezoelectric accelerometers is associated with their advantage over other known types of accelerometers. Piezoelectric accelerometers have a wide operating frequency range, active conversion, have a linear amplitude characteristic, high resistance to external influences. The absence of moving parts in them ensures high reliability and durability; they can be implemented in a small-sized design with high manufacturability in production.

In known accelerometers, the sensitivity vector and direction of measured inertial force have opposite directions, which makes it possible to measure only linear inertia and acceleration within small limits. Such accelerometers do not allow measuring the speeds of moving objects in real time. Due to the presence of elements that create electromagnetic fields, their design becomes more complicated and requires additional measures to exclude the influence of these electromagnetic fields [7-10].

In this regard, it becomes necessary to develop a three-dimensional piezoelectric accelerometer for measuring the dynamic parameters of moving objects, a mathematical basis for the construction and a mathematical model of its functioning, in which the polarization vector of the piezoelectric element is directed along the axis of sensitivity of the accelerometer, i.e. direction of the vector of linear acceleration of moving objects. Along with this, below is the classification of accelerometers, their characteristics and parameters measured by the accelerometer, objective function is formulated.

Development of target function

To assess the effectiveness of a three-dimensional piezoelectric accelerometer for measuring the dynamic parameters of moving objects, determined by an increase in accuracy and reliability of measurements, sensitivity and functionality expansion of this accelerometer with a rapidly changing direction of object measurement movement along the three axes "XYZ" of a rectangular coordinate system, an objective function is formulated in an automatic mode. Accuracy of the accelerometer depends on sensitivity of sensors and dynamic range, and sensitivity, i.e. amplitude with which the device emits an electrical signal, which is considered main parameter when choosing a measuring device. The accuracy of data obtained during the measurement depends on it. The dynamic range of vibrations captured by accelerometer is determined by vibration energy, which is in a very narrow frequency range. Therefore, it is necessary to take into account that the vibration measurement process is performed with very small or very large amplitudes of acceleration and, accordingly, speed.

Improving the efficiency of this accelerometer operation consists in solving two main tasks, i.e. in the development of objective function and determination of optimal values of its arguments, which would ensure the operation of accelerometer with a rapidly changing direction of movement of measurement object along the three axes "XYZ" of a rectangular coordinate system in automatic mode.

Guided by the above principles, the target function of a three-dimensional piezoelectric accelerometer can be determined in the following form:

$$E_{TPA} = \{\min[t_{iz}, \delta_{iz}, P_{pom}], \max[\tau_{l_1}, A_{ch}, D_{iz}]\},$$

where t_{iz} – measurement time; δ_{iz} – measurement error; P_{pom} – noise level along two other coordinates, along which measurements are not being performed at a given time; τ_{1_1} – accelerometer speed; A_{ch} – the sensitivity of the accelerometer, i.e. amplitude, starting from which the accelerometer emits an electrical signal and is considered the main parameter when choosing a measuring device; D_{iz} – dynamic range of measurement of an accelerometer, which is determined by the upper and lower values of vibrations that the accelerometer is capable of detecting and is in a very narrow frequency range.

Development of a three-dimensional piezoelectric accelerometer for measuring the dynamic parameters of moving objects

To expand the functionality, increase the accuracy, sensitivity and reliability of measuring dynamic parameters of moving objects relative to the coordinate axes "XYZ" of a rectangular coordinate system in the automatic measurement mode with a rapidly changing direction of movement of the measurement object, a design scheme for the accelerometer for measuring the dynamic parameters of moving objects in automatic mode has been developed "and its control scheme, which is shown in Fig. 1 [1, 4, 5].

The developed accelerometer contains a piezoelectric element along each coordinate, made in the form of a two-layer flat plate, glued or welded together by wide edges having an excitation section – 1,2,3 and a sensitive element section – 4,5,6, isolated from each other by insulating grooves – 7,8,9, and the piezoelectric elements are damped relative to the longitudinal coordinate axes through "P" – shaped flat springs – 10,11,12, while the measurement object – 13 is equipped with a three-coordinate position sensor – 14, the electrodes – 15 of which are connected to the first input of the first differential operational amplifier – 16, to the second input of which the output of the sampling unit is connected – 17 axes of the rectangular coordinate system "XYZ", the first – 18, the second – 19, the third – 20, the outputs of the first differential operational amplifier along three coordinate axes are connected to the inputs of the threshold device – 21, which, depending on the set operation threshold does not pass the generated interference signals in two other coordinates to the output, the recognition unit – 22 coordinate signals with the ability to process a random sequence, the outputs of the threshold device are connected to the inputs of the memory unit – 23, which is connected to the control code generator – 24 and to the control unit – 25, the outputs of which along the three axes "XYZ" are connected to the control inputs of the triggers – 26,27,28, выходы которых соединены к входам электронных ключей – 29,30,31, signal inputs and outputs of which are respectively connected to the inputs of the alternating voltage source – 32,33,34 coordinate axes "XYZ" and to the excitation electrodes of piezoelectric elements, and the electrodes of the sensitive elements of the coordinate axes "XYZ" are connected to the inputs of differential operational amplifiers – 35.36, 37 coordinate axes "XYZ", the first outputs of which through pulse shapers – 38,39,40 are connected to the first digital indicators – 41,42,43, and the second outputs of differential operational amplifiers through pulse shapers – 44,45,46 and integrators – 47,48,49 are connected to the inputs of the second digital indicators - 50.51.52. Inertial masses – 53,54,55 are mechanically connected with piezoelectric sensing elements, piezoelectric elements are installed on "P" -shaped flat springs. Each of the piezoelectric sensing elements, located along the XYZ axes, together with the mechanically connected inertial mass, is placed inside a separate protective casing – 56,57,58.

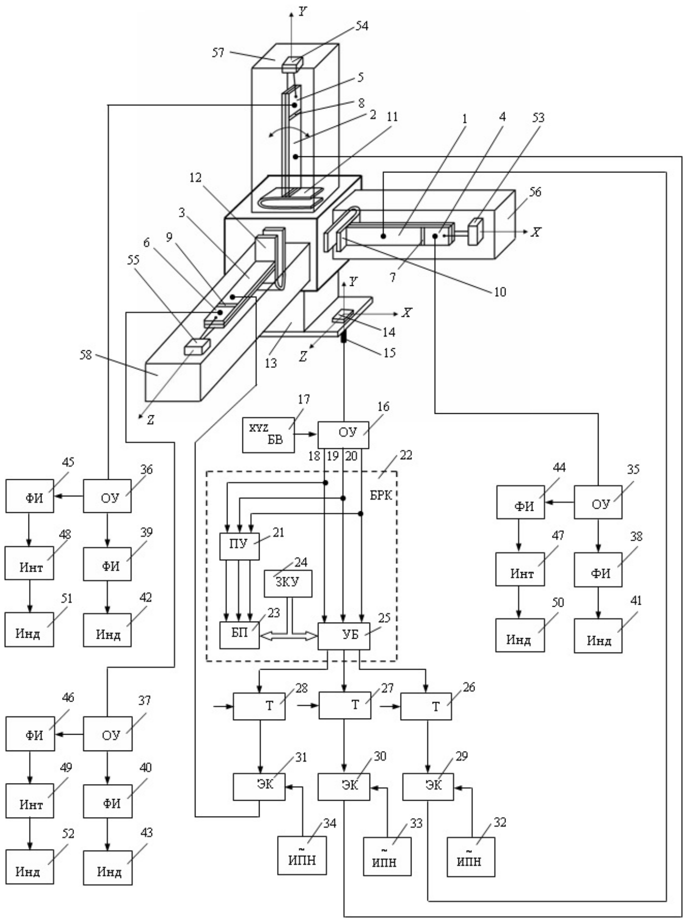

Figure 1. Diagram of the design of a three-dimensional piezoelectric accelerometer for measuring the dynamic parameters of moving measurement objects with a control circuit

Operating principle

The principle of the accelerometer operation is based on the fact that, depending on required direction of measurement object movement – 13 at the beginning in the sampling block – 17, the initial state is set along one of the XYZ axes of a rectangular coordinate system [5, 6].

When the measurement object moves – 13 V along the "X" axis, first the three-coordinate position sensor - 14 perceives the movement of the measurement object - 13 and signals from its output are fed to the input of the first differential operational amplifier – 16. Amplified signals, passing through the threshold device – 21, which depending on the set triggering threshold, it does not pass the generated interference signals along two other coordinates to the output, they enter the coordinate recognition unit – 22, which are fed into the memory unit – 23. The signals from the output of control code generator – 24 are fed to the input of memory unit – 23 and the controlled unit – 25, from output of which the signals are recognized by their amplitude, by their frequency or by their average value in allocated frequency band. Since the maximum signal amplitude occurs along the "Y" axis, since the direction of sensitivity axis in three-coordinate sensor – 14 along the "Y" axis of rectangular coordinate system coincides with direction of measurement object movement – 13 along the "X" axis, signals from the output of coordinate recognition unit – 22 axis "Y" is fed to the trigger control input – 27, from the output of which signals sent to the control input of electronic key – 30. In this case, the signals from output of AC voltage source – 33 are fed to the signal input of electronic switch – 30, from the signal output of which voltage is fed to the electrodes of excitation section – 2 piezoelectric elements, which create alternating bending vibrations in it, the direction of which coincides with the direction of sensitivity axis vector of section sensing element – 5.

With the linear movement of the measurement object – 3, depending on the change in speed in the direction along the "X" axis, the inertial mass – 54, due to its inertia, lags behind or ahead of the body of the measurement object – 13, which creates the effect of inertial forces on the sensitive element – 5, leading it to bending deformations that generate signals proportional to the varying acceleration. These signals are fed to a differential operational amplifier – 36.

Strengthening, these signals through the pulse shaper – 39 are fed to the digital indicator – 42, showing the measured acceleration of the linear motion of the measurement object – 13, and the signals passing through the pulse shaper – 45, the integrator - 48 and the digital indicator – 51 shows the measured speed of the measurement object – 13.

The measurement of the dynamic parameters of the movement of the measurement object – 13 relative to the other two axes "Y" and "Z" of the rectangular coordinate system is carried out in an automatic mode of operation similarly in the above sequence. With the help of a three-coordinate position sensor – 14, a differential operational amplifier – 16 and a coordinate recognition unit – 22, automatic switching of the coordinate axes "XYZ" is carried out depending on the change in the direction of movement of the measured object – 13, inertial masses – 53,54,55 are mechanically connected with piezoelectric sensitive elements. Each of the piezoelectric sensing elements, located along the XYZ axes, together with an inertial mass mechanically connected to it, is located inside a separate protective casing – 56,57,58.

Mathematical basis of the accelerometer functioning

In the developed accelerometer, it is proposed to use a piezoelectric element as a working element, made in the form of two-layer flat plates along each coordinate axis X, Y and Z, glued or welded together by wide edges having an excitation section and a sensitive element section, depending on the change in direction of movement measurement object.

The use of a piezoelectric element as a working (converting) organ is associated with its advantages listed above.

When the measurement object moves along one of the coordinate axis, an effect is created on sensitivity sections of accelerometer piezoelectric element, measured in units of inertial gravity, and as a result, the voltage components obtained from electrodes of piezoelectric element are created.

The moment of piezoelectric element inertia is associated with geometric shape of inertial mass and is determined as follows [4-6]:

$$i = l_2 \cdot l_3^3 / 12, (\text{m}^4) \tag{1}$$

where l_2 – piezoelectric element width (m); l_3 – piezoelectric element thickness (m).

The stiffness of accelerometer piezoelectric element is determined by following formula:

$$C = l_1 / 3 \cdot E_{ю} \cdot i, (\text{N/m}) \tag{2}$$

where l_1 – length of piezoelectric element (m); $E_{ю}$ – Young's modulus (modulus of elasticity) (N/m^2); i – moment of inertia of the piezoelectric element (m^4).

According to expression (2) in known accelerometer, the piezoelectric element has a relatively high stiffness value [1] and therefore, to remove a signal with a voltage of 0.5 mV from the electrodes of a sensitive piezoelectric element, the value of the linear inertial force should be 5-6 times higher than the value of this force, acting on a sensitive piezoelectric element. Since the value of the linear inertial force is at a low sensitive area, it allows you to measure relatively smaller values of this force.

By differentiating the velocity vector of free end of piezoelectric element by time vector, one can determine the acceleration of free end of piezoelectric element by following formula:

$$\vec{i} = \frac{d\vec{v}}{dt} = \frac{d^2 r}{dt}, \tag{3}$$

where $\vec{v}$ – piezoelectric element free end speed (m/s); t – time (s); r – trajectory of movement.

If we take into account the trajectories of measurement object along one of the coordinate axis, then at some point in time $-t_0$ trajectory vector $\vec{r}(t) = \vec{r}_0$ and time dependence of acceleration vector of free end of piezoelectric element $\vec{i} = \varphi(t)$.

Then integrating formulas (3) on the time interval from t_0 to t we obtain the following dependences to determine $\vec{v}$ – the speed of free end of piezoelectric element and $\vec{r}(t)$ – motion path vector:

$$\vec{v}(t) = \vec{v}_0 + \int_{t_0}^{t} (t)dt, \tag{4}$$

$$\vec{r}(t) = \vec{r}_0 + (t - t_0)\vec{v}_0 + \int_{t_0}^{t}\int_{t_0}^{t} \vec{i}\, dt^2. \quad (5)$$

If the time vector $\vec{i}$ does not change, then the given motion of the measurement object is considered uniformly accelerated and, transforming formulas (4) and (5), we obtain expressions for determining the speed of the free end of the piezoelectric element and the vector of the motion trajectory:

$$\vec{v}(t) = \vec{v}_0 + (t - t_0) \cdot \vec{i}, \quad (6)$$

$$\vec{r}(t) = \vec{r} + (t - t_0) \cdot \vec{v}_0 + \frac{(t - t_0)^2}{2} \cdot \vec{i}. \quad (7)$$

Depending on the purpose of accelerometer, mathematical basis makes it possible to determine necessary nodes when building a control circuit for a three-coordinate piezoelectric accelerometer for measuring dynamic parameters of moving objects in an automatic mode.

Mathematical model of functioning

Mathematical model is built according to the following sequence. First, to calculate the main parameters of piezoelectric element, we determinedisplacement amplitude of free end of piezoelectric element, made in form of two-layer flat plates with a length l_1 with longitudinal vibrations of first mode along its length according to following formula:

$$\Delta_{PA} = \frac{2T_c \cdot l_1}{\pi \cdot E_u} = \frac{2E_{\sim} \cdot d_{31} \cdot Q_m \cdot E_u \cdot l_1}{\pi}, \text{ m} \quad (8)$$

where T_c – maximum allowable mechanical stress (N/m^2).

The maximum allowable mechanical stress arising in a deformable piezoelectric element during the movement of the measurement object is determined by following expression:

$$T_c = E_{\sim} d_{31} \cdot Q_m \cdot E_u, \text{ N/m}^2 \quad (9)$$

where $E_{\sim}$ – voltage of electric field applied to the plates of piezoelectric element made in form of two-layer flat plates (V/m); d_{31} – piezoelectric module for piezoceramics TsTBS-3 (piezoelectric ceramics from lead zirconate-titanate) (m/V); Q_m – mechanical quality factor, determined experimentally from amplitude-frequency (resonance) characteristics of piezoelectric element (Fig. 2); E_u – Young's modulus for a piezoelectric element of the TsTBS-3 brand (N/m^2).

If we do not take into account the structural losses, then we can obtain dependence for maximum permissible linear velocity of free end of piezoelectric element:

$$V_{pr.dop} = \frac{4F_c \cdot V_{zv}}{\pi \cdot l_1 \cdot l_2 \cos\alpha \cdot E_u} = \frac{4F_c}{\pi \cos\alpha \cdot Z_0} = \frac{4F_c V_{zv}}{\pi \cos\alpha \cdot E_u}, \text{ (m/s)} \quad (10)$$

where $Z_0 = l_1 \cdot l_2 \cdot \sqrt{\rho \cdot E_u}$ – wave resistance; density of CBTS-3 material (kg/m^3).

The resonant frequency of a piezoelectric element made in the form of two-layer flat plates can be determined by following formula:

$$f_p = V_{zv} / 2 \cdot l, \text{ (kHz)} \quad (11)$$

where $V_{зв}$ – longitudinal wave propagation speed (m/s).

The speed of the working end of a loaded piezoelectric element made in form of two-layer flat plates is determined as follows:

$$V_{pkn} = \frac{U_{\sim vozb} \cdot l_2 \cdot d_{31} \cdot E_u}{R_{mp} \cdot \cos\alpha}, \text{ (m/s)} \quad (12)$$

where α – the angle of inclination of piezoelectric element made in form of two-layer flat plates to working body (rad); resistance to mechanical losses.

Mechanical loss resistance is defined as follows:

$$R_{MP} = \omega \cdot M / Q_m, \text{ (N s/m)} \quad (13)$$

The average linear speed of working body movement is determined by following formula:

$$V_{po} = \Delta_{PA} \cos\alpha \cdot k. \text{ (m/s)} \quad (14)$$

Taking into account the complex linear nature of friction value, correction factor for average linear speed of movement of working body is determined by formula:

$$k = V_{po} / V_{pkn}, \quad (15)$$

where V_{po} – linear speed of working body movement (m/s); $V_{ркн}$ – speed of the end of working body of the loaded piezoelectric element made in form of two-layer flat plates (m/s).

Taking into account the losses associated with the action of a constant clamping force of the piezoelectric element on working element, speed of the end of working element of loaded piezoelectric element is determined as follows:

$$V_{pk} = V_{pkn} / 2, \text{ (m/s).} \quad (16)$$

The nominal moment acting on piezoelectric element is determined by following formula:

$$M_H = F_T \cdot R, \text{ (Nm)} \quad (17)$$

where F_T – maximum tangential force on the piezoelectric element (N); R – bending radius of the piezoelectric motor (m).

The maximum tangential force on a piezoelectric element is determined by following expression:

$$F_T = U_{vozb} \cdot l_2 \cdot d_{31} \cdot E_u \cos\alpha. \text{ (H)} \qquad (18)$$

where U_{vozb} – AC source voltage (V); l_2 – piezoelectric element thickness (m); d_{31} – piezoelectric module of piezoceramics (m/V); E_u – Young's modulus (modulus of longitudinal elasticity), i.e. physical quantity characterizing the ability of a material to resist stretching, compression in bending (N/m^2).

Taking into account equation (18) in equation (17), we obtain the following expression for a piezoelectric element made in form of two-layer flat plates glued or welded together by wide edges having an excitation section and a sensitive element section:

- - for a piezoelectric element made in the form of two-layer flat plates:

$$M_1 = U_{vozb} \cdot l_2 \cdot d_{31} \cdot E_u \cos\alpha \cdot R, \text{ (Nm)} \qquad (19)$$

- - for "P" -shaped flat springs, with which the piezoelectric elements are damped relative to the longitudinal coordinate axes:

$$M_2 = F_{pr} \cdot R, \text{ (Nm)} \qquad (20)$$

where F_{pr} – clamping force of "P"-shaped flat springs (N).

The moment created by the tangential force and the force of "P" -shaped flat springs acting on a piezoelectric element made in form of two-layer flat plates is determined as follows:

$$M = M_1 + M_2, \text{ (Nm).} \qquad (21)$$

The time required for starting and establishing mechanical vibrations of a piezoelectric element made in form of two-layer flat plates is determined as follows:

$$t_{пуск} = 1{,}47 \cdot \frac{Q_м}{f_p}, \text{ (c).} \qquad (22)$$

where f_p – resonant frequency of the supply voltage (Hz); $Q_м$ – mechanical figure of merit of piezoelectric element.

The mechanical figure of piezoelectric element merit made in form of two-layer flat plates is determined from the amplitude-frequency characteristic (fig. 2) according to following formula:

$$Q_m = f_p / \Delta f, \qquad (23)$$

where Δf – bandwidth; f_p – mechanical resonance frequency (Hz).

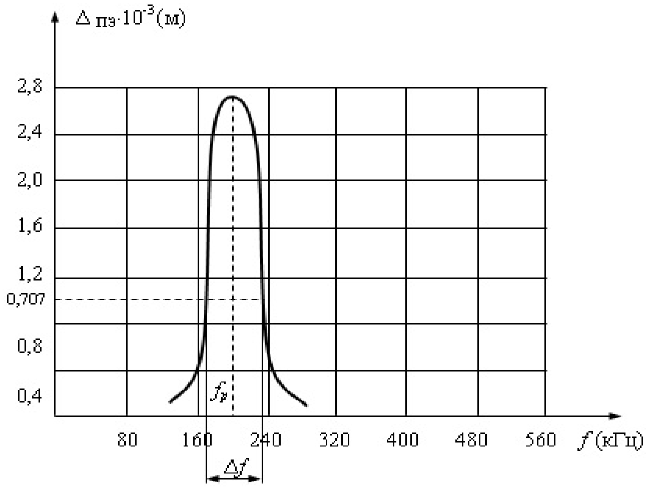

Figure 2. Amplitude-frequency (resonance) characteristic of piezoelectric element

The speed of accelerometer depends on frequency of mechanical resonance of piezoelectric element made in form of two-layer flat plates, parameters of electrical circuit of power supply and is determined by following formula:

$$\tau_{l_1} = 1/2 \cdot f_p. \text{ (Hz)} \qquad (24)$$

As follows from the above analytical dependences, which were used to determine the main parameters of accelerometer, coincide with experimental data and discrepancy between them is 1-3%.

Conclusion

Thus, the developed three-coordinate piezoelectric accelerometer for measuring the dynamic parameters of measurement moving objects in automatic mode has the following advantages over known accelerometers:

1. 1. In the proposed accelerometer for measuring the dynamic parameters of a moving measurement object in automatic mode, the supply of a position sensor of measurement object using a signal recognition unit of three coordinate axes "XYZ" of a rectangular coordinate system ensures determination of movement direction, i.e. coordinate axis X, Y or Z with measurement of acceleration or linear velocity when measurement moving object along this coordinate axis.
2. 2. In the proposed accelerometer, connection of second outputs of differential operational amplifiers through pulse shapers and integrators connected to inputs of second digital indicators provides measurement of movement speed for measurement object relative to three axes "XYZ" of rectangular coordinate system.
3. 3. In the proposed accelerometer, due to the damping of piezoelectric elements with the help of "P" -shaped flat springs along three axes "XYZ" of trectangular coordinate system when e measurement object moves along one of the coordinate axes, along other two axes, generated noise signals in amplitude are significantly reduced by their suppression, since the directions of mechanical vibrations that create amplitudes

of interference are located along the axes of rectangular coordinate system of piezoelectric elements.

4. 4. In the proposed accelerometer, the presence in construction of piezoelectric element of a section for excitation of mechanical vibrations, as well as a section of a sensitive element that generates signals with rapidly changing acceleration and speed of a moving object of measurement, the signal amplitude and sensitivity increase 4-5 times due to the generation of a signal by sensitive elements in vibration excitation mode created in the accelerometer.

Comparative analysis shows that the above advantages of developed accelerometer directly affect the achievement of specified technical result, i.e. allow to expand functionality, increase the sensitivity and reliability of process of measuring dynamic parameters of a moving object in measurement relative to the "XYZ" axes of a rectangular coordinate system in automatic mode with a rapidly changing direction of measurement object movement.

References

1. A.A. Brickichev. Piezoelectric accelerometer. Patent of the Russian Federation No. 2400760. http://www.findpatent.ru/patent/240/2400760.html © FindPatent .ru - patent search, 2012 - 2015.
2. V.V. Jancic. Piezoelectric vibration measuring converters (accelerometers). Rostov-on-Don: News of the Southern Federal University, 2010. 304 p.
3. V.M. Sharapov, M.P. Musienko, E.V. Sharapova Piezoelectric sensors. Ed. V.M. Sharapova. Moscow: Technosphere, 2006. 632 p.
4. T.M. Mansurov, Ch.P. Aliyev. Three-coordinate accelerometer for measurement of acceleration and speed of measurement object in automatic mode. *Bulletin of the Eurasian Patent Office (EAPV) of the Eurasian Patent Organization* (priority of application № 201700327/31). 2018. No. 9. P. 125. 28.09.2018.
5. T.M. Mansurov, Ch.P. Aliyev. Accelerometer for measurement of accelerations and speeds of moving objects in automatic mode. *Journal of Khmelnitskii National University*. Kharkov, CNU. No. 5, 2018. P. 78-83.
6. Yu.I. Iorish. Vibrometriya. Second edition, revised and supplemented. Moscow: State Scientific and Technical Publishing House of Machine-Building Literature, 1963. 771 p.
7. M.D. Genkin, A.G. Sokolova. Vibrations in Engineering. Moscow: Mechanical engineering, 1981. Vol. 5. Measurements and tests. Ed. M.D. Genkin. 1981. 496 p.
8. K. Brindley. Transmitters: Reference: Per. Moscow: Energoatomizdat, 1991. 144 p.
9. G. Vigleb. Sensors. Device and application. Moscow: World, 1989. 196 p.

METHODS FOR DESIGNING PROFESSIONAL COMMUNICATION SYSTEMS OF TETRA STANDARD, TAKING INTO ACCOUNT RELIABILITY INDICATORS

Mohammed Omar Ahmed Abdulvasea,
Sana'a University, Sana'a, Republic of Yemen

Kulpunai A. Adylbekova,
Kyrgyz State Technical University, Bishkek, Kyrgyz Republic

DOI: 10.36724/2664-066X-2020-6-6-17-20

ABSTRACT

The design of professional radiotelephone communication systems within framework of graduation qualification works (WRC) is an important part of educational process, which allows you to apply theoretical knowledge on use of equipment of current TETRA standard for deployment of alternative communication systems. And also to gain skills in applying methods of teletraffic theory and reliability theory for design and solution of practical problems. It is proposed to design a TETRA standard system for mastering the competencies of an educational standard in two stages. Namely, to carry out preliminary planning, in which the required number of radio channels and the range of action are calculated with engineering precision, as well as a preliminary estimate of the amount of equipment required for the deployment of the system. Next, an assessment is made ofreliability indicators impact on the quality of the system, namely, on probability of call losses in order to substantiate necessary reserve and select option for organizing work.using SIP protocols are relevant.

KEYWORDS: *Unified dispatching services, design, communication modes, service disciplines, likelihood of maintaining operability, likelihood of call losses.*

INTRODUCTION

A feature of so-called trunking systems of service radiotelephone communication is operation in a dedicated frequency range that differs from the frequency ranges of public mobile communication systems. In case of emergencies, as well as during public events, professional mobile communication becomes an alternative to overloaded systems of GSM and LTE standards [1,2,3].

Communication equipment manufacturers offer a wide range of professional radiotelephone communication systems - analog and digital. However, users are focused on the transition to digital standards. In this regard, it is advisable to carry out the implementation of promising programs for the modernization of professional radio communication networks, relying on modern digital trunking communication systems and, in particular, on the digital TETRA standard systems.

To access the specifications of the open standard TETRA communication, you must register as a member of the Association "Memorandum of Understanding and Promotion of TETRA Standard". The equipment of base and user stations of TETRA standard is offered by a number of companies - system integrators. A feature of TETRA standard systems is the ability to form complex network structures based on fiber-otic rings with a sufficient distance of transceivers from each other. It becomes possible to organize work at the same frequencies, using them repeatedly.

In recent years, dozens projects of TETRA standard systems of various scales have been implemented – for transport and industrial enterprises, sports facilities, oil and gas industry and energy enterprises. No license is required to deploy trunking systems. We can mention the largest project – the integration into a single TETRA network of systems built within the framework of the projects of the oil pipelines "Eastern Siberia – Pacific Ocean", "PS-21 ESPO – border of the PRC" and "Anzhero-Sudzhensk – Taishet" by JSC "Transneft". In fact, we are talking about the deployment of technological communication systems of considerable length.

Users are offered special means of protection against unauthorized access and a number of additional services, the implementation of which is relevant for law enforcement agencies and rescue services. More than 6 million subscribers use service radiotelephone communication systems around the world.

Taking into account the increased terrorist activity around the world, it is important to combine and promptly process emergency traffic in unified dispatch services (UDDS). In this issue, a new technological level of access means for professional radiotelephone communication becomes important. The experience of implementing EDDS in the Republic of Yemen is at the stage of pilot projects. It is necessary to develop methods for dynamic management of UDDS resources as a result of the development of generalized criteria for assessing communication quality indicators, taking into account reliability indicators.

The TETRA standard radio interface provides for operation in a frequency grid with an interval of 25 kHz. The minimum duplex distance between duplex radio channels is 10 MHz. Special frequency bands may be allocated for TETRA systems. So, in European countries, the frequency ranges 380-385/390-395 MHz are assigned to the security services, and 410-430/450-470 MHz ranges are provided for commercial organizations. In Asia, the 806-870 MHz band is used for TETRA systems. The operating frequency range of TETRA system in the Russian Federation is 410-430/450-470 MHz.

Initially, trunking mobile communication systems were created on the assumption that almost all traffic would be locked inside the communication system. The economical use of the radio resource was achieved by using half-duplex communication, which assumes alternate use of the radio channel by users. Both radio stations automatically retune to the transmission Ftrans and the reception Fpr, which form one traffic channel. Subscribers start negotiations using these frequencies in turn. When any subscriber presses the "hang up" key, the radio stations automatically return to standby mode on the control channel. Thus, to establish a connection between two mobile subscribers of a trunking system, one traffic channel is required. The half-duplex principle is basis of low-cost networks that connect dozens of subscribers in various parts of city and in open areas.

In digital trunking systems and, in particular, in TETRA standard systems, it is possible to use a duplex communication option between two users of TETRA system. Duplex communication employs two frequencies. This allows for a familiar dialogue. When organizing the interaction of the trunking system with public switched telephone network (PSTN), the Ministry of Communications and Communications of the Russian Federation recommended duplex operation.

The communication system of TETRA standard is proposed to be considered as a queuing system (QS), using the methods of teletraffic theory for its description, which establish the relationship between the amount and nature of traffic load, number of serving devices (number of temporary traffic channels in radio interface) and quality of service [4]. The system of TETRA standard contains all characteristics of QS necessary for this: simplest flow of applications; duration of a radio channel being occupied by a call; number of channels provided to users.

It is possible to use one of three disciplines of call handling:

– with refusa s (call receives a refusal of service if there are no free servicing devices at the time of the call);

– with waiting (call enters the queue and waits for device to be released if there is no free serving device at the time of call);

– combined service discipline (waiting time limits are imposed - by waiting time or by the length of queue).

A feature of the systems of TETRA standard is the ability to establish these disciplines in accordance with the selected comparison criterion.

Traditionally, such a criterion is the probability of denial of service calls in a service with failures, or the probability of waiting for start of service of a call over a given time in a waiting service.

The estimated subscriber capacity of the system affects the required number of radio channels and, accordingly, the number of transceivers. Figure 1 shows the calculations of temporary channels required number in radio interface V depending on expected capacity of communication system N_Σ in range N_Σ= 500…2000 subscribers. The calculations were carried out for a TETRA communication system with a "star" configuration with a compact arrangement of transceivers:

– for service with waiting at a fixed probability of waiting for the start of service $P(t_o > \tau)$ = const, where t_o – service start waiting time, relative value τ defined as $\tau = t_o/t_{ob}$;

– for service with failures at a given probability of call loss P_p = const.

A feature of TETRA system is possibility of a gradual failure of equipment with possibility of degrading the quality of service. Let us calculate the possible increase in probability of denial of service for a discipline with refusals.

The calculation was carried out using the first Erlang formula. It is assumed that the number of traffic channels of one base station can abruptly decrease from the initial value V = 4 x 4 = 16 traffic channels (there are four transceivers in operation, each of which provides a fourfold time division multiplexing of radio channels). In case of failure of one transceiver, the number of traffic channels of the base station will be V = 4 x 3 = 12 traffic channels. The calculation was also carried out for V = 8 and 4 traffic channels available to users. A complete failure occurs if the failure rate exceeds 5 -10% or service is stopped altogether. The calculation results are presented in table 1.

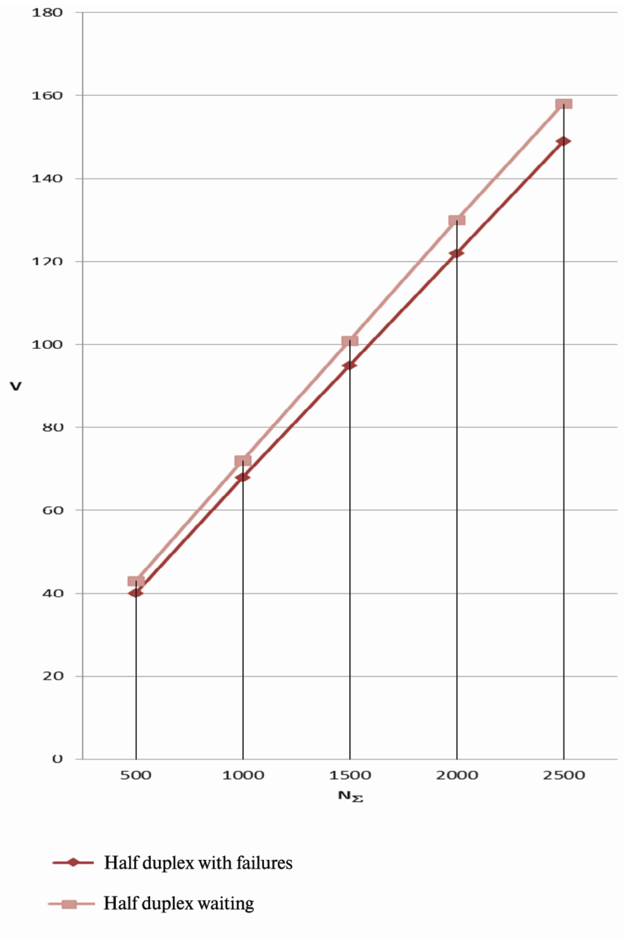

Figure 1. Dependence of the number of traffic radio channels V on the number of subscribers of the system N_Σ, for half-duplex communication with failures and with waiting

Table 1

Calculation results of the probability of call loss under conditions of gradual failure of receiving-transmitting equipment for Y = 5 Earl in duplex mode

Call loss probability, P	**Number of available traffic channels**			
	V=16	V	V=8	V=4
	0,000049	0,003441	0,070048	0,398343

From the point of view of the theory of reliability, TETRA equipment can be considered as a system with loaded, unloaded or sliding redundancy, for the calculation of which the formulas of serial and parallel connection of elements can be used. In this case, it is considered that the backup elements operate in the main mode both before and after their failure, therefore, reliability of backup elements does not depend on moment of their transition from the standby state to main one and is equal of main elements reliability.

In case of unloaded redundancy, the backup elements are sequentially switched on in case of failure of the main and then the first backup element, therefore, the reliability of backup elements depends on moment of their transition to the main state. Such redundancy occurs most frequently in various technical systems, since it is similar to replacing failed elements and assemblies with spare ones.

A system of “m out of n” type can be considered as a variant of a system with parallel connection of elements, the failure of which will occur if less than m elements (m <n) out of n elements connected in parallel are operable.

При m=1 система превращается в обычную систему с параллельным соединением элементов, а при m=n – с последовательным соединением.

Table 2 shows formulas for calculating the probability of failure-free operation of systems of the "m out of n" type at m≤n≤5 [5].

Table 2

Probabilities of failure-free operation system ("m of n")

m	Total number of elements, n				
	1	2	3	4	5
1	p	$2p-p^2$	$3p-3p^2+p^3$	$4p-6p^2+4p^3-p^4$	$5p-10p^2+10p^3-5p^4+p^5$
2	-	p^2	$3p^2-2p^3$	$6p^2-8p^3+3p^4$	$10p^2-20p^3+15p^4-4p^5$
3	-	-	p^3	$4p^3-3p^4$	$10p^3-15p^4+6p^5$
4	-	-	-	P^4	$5p^4-4p^5$
5	-	-	-	-	p^5

It can be concluded that in a specific example, a sharp decrease in the quality of service, expressed in an increase in number of failures, occurs if only two of four transceivers remain in operation, and in the event of failure of three of four transceivers, a non-stationary mode of operation occurs (load volume exceeds the capabilities of the system). It is proposed to consider from the point of view of reliability the TETRA base station of the UDDS office as a system of the "2 out of 4" type, for the description of which an expression from Table 1 of the form

$$P = 6p^2 - 8p^3 + 3p^4, \tag{1}$$

where P is the probability of keeping the system working; p is the probability of finding one transceiver of the system in a working state.

If we assume that the value of p = 0.95, then the value of P will be equal to

$$P = 6 \text{ x } (0{,}95)^2 - 8 \text{ x } (0{,}95)p^3 + 3 \text{ x } (0{,}95)^4 =$$
$$= 0{,}9995186.$$

If we assume that the value of p = 0.99, then the value of P will be equal to

$$P = 6 \text{ x } (0{,}99)^2 - 8 \text{ x } (0{,}99)p^3 + 3 \text{ x } (0{,}99)^4 = 0{,}999996.$$

Figure 2 shows a specific TETRA communication system presented in terms of reliability. We have two base stations operating in load sharing mode. They are described as a system of two parallel elements. Each such element contains sequentially connected elements – a control unit (CU) and transceiving equipment as a "2 of 4" scheme.

Let's imagine the first base station as a certain element A, describe its performance in the form of a sequential diagram as

$$P_{БС} = P_A = P_{УУ} \times P_{«2 из4»}, \quad (2)$$

where P_{BS} – probability of maintaining operability of base station; P_{CD} – probability of maintaining operability of control device; $R_{"2 \text{ out of } 4"}$ – probability of maintaining operability of transmitting and receiving equipment.

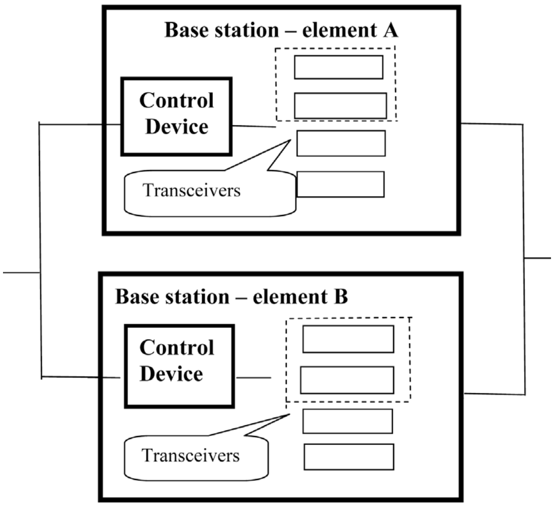

Figure 2. Representation of a back office TETRA system as an equivalent circuit for assessing reliability

The probability of maintaining the operability of entire TETRA system of UDDS office, calculated as a parallel circuit of connecting two base stations, will be

$$P_{TETRA} = 1 - (1 - P_{BS})^2. \quad (3)$$

Assuming that $P_{«2 \text{ from } 4»} = 0{,}9995186$ and $P_{УУ} = 0{,}999$, we get

$$P_{БС} = 0{,}9995186 \times 0{,}999 = 0{,}998519.$$

Then we will continue the calculation by formula (3) and get

$$P_{TETRA} = 1 - (1 - 0{,}998519)^2 = 0{,}9999979.$$

Thus, the proposed service organization is highly reliable due to use of two base stations. The highest reliability requirements should be placed on the control device of each base station.

REFERENCES

1. http://ess.ru. Comparative analysis of digital trunking communication standards.
2. http://citforum.ru. TETRA professional radio standard. Benefits and opportunities.
3. Motorola equipment for TETRA systems [Electronic resource] http://www.vseradio.ru/catalog/radiostations/canopy_tetra/motorola_tetra.php, 2015.
4. B.S. Livshits A.P. Pshenichnikov, A.D. Kharkevich. Teletraffic theory. Moscow: Communication, 1979. 224 p.
5. I.A. Ushakov. System reliability theory course. Moscow: Bustard, 2008. 240 p.

INCREASING THE CAPACITY OF FIBER-OPTICAL TRANSMISSION SYSTEMS DUE TO DECREASING DISTANCES BETWEEN BEARING

Augustin Vyukusenge,
University of Bujumbura, Bujumbura, Burundi,
vyukusengeaugustin@yahoo.fr

DOI: 10.36724/2664-066X-2020-6-6-21-23

ABSTRACT

The load on transport networks based on fiber-optic transmission systems is increasing at an accelerating rate. This paper discusses the possibility and limitations of increasing the throughput of fiber-optic transmission systems by reducing the distance between carriers. A comparison is made between fixed and flexible grids in terms of the spectral bandwidth efficiency. It is concluded that the use of flexible mesh technology is promising when switching to channel speeds above 100 Gbit / s.

KEYWORDS: *Bandwidth, fixed mesh, flexible mesh, number of channels, spectral efficiency, fiber optic transmission system.*

INTRODUCTION

The current period of development of the information and communication technologies (ICT) sector is characterized by the widespread introduction of a variety of multimedia services, the rapid development of Internet networks, the emergence of data processing centers, the massive introduction of mobile applications requiring high bandwidth and advanced technologies for high-speed packet transmission. The implementation of these trends necessitates a sharp increase in the throughput of fiber-optic transmission systems (FOTS). Obviously, under these conditions, a transmission medium is of particular interest, which has a high potential throughput and allows the volume of transmitted information to be multiplied. Thus, fiber-optic cables act as the main transmission medium of the transport network, on the basis of which a layer of transparent optical channels is formed using the technology of spectral multiplexing (DWDM). The data transfer rate achieved by fiber-optic systems over the past 30 years has increased by more than four orders of magnitude [1]. Traffic forecasting in the construction of optical communication networks is a difficult task, which requires the development of methods to increase the bandwidth of FOTS without significant costs for the modernization of expensive line-cable facilities. In this regard, in order to expand the physical network, it is advisable to make more significant efforts to improve the efficiency of fiber-optic information transmission systems. One of the ways to effectively use the capabilities of optical fiber is to increase the number of channels in a frequency band by reducing the distance between channels, the path along which the developers of the G.692 standard [7] went.

Applying a fixed grid

The number of channels, channel spacing, width of each channel and channel bandwidth are important parameters in design and construction of a high-speed communication network.

In traditional DWDM systems, the optical spectrum in C-band, consisting of approximately 4.1 THz, is divided into hard spectral intervals of 50 GHz as defined in ITU-T Rec. G.694.1. This forms a "grid of wavelengths, where the center frequencies of adjacent channels have fixed spectral intervals of 50 GHz. The frequency grid defined in this recommendation supports a variety of fixed channel spacings from 12.5 GHz to 100 GHz or more (integer multiples of 100 GHz), as well as flexible grid [2].

When using DWDM technology, one can try to estimate the limiting value of equivalent FOTS bandwidth, a parameter defined as the product of transmission rate in optical channel B_{ch} for the number of channels N.

$$B_{max} = B_{ch} \times N ,$$

Using the standardized range (192.10 - 196.10 THz), the total frequency width is 4.1 THz. If the channel spacing is 50 GHz, a maximum of 81 channels can be accommodated. If the channel spacing is reduced to 25 GHz, this range can accommodate 163 channels, and thus double the bandwidth of the FOTS.

Although systems with reduced channel spacing will be able to provide significant FOTS bandwidth, this decrease places more stringent demands on the devices used in the system, which reduces the number of potential equipment manufacturers and also increases its cost. From the transmitter's point of view, wavelength stability becomes very important, since even a small drift can cause serious inter-channel interference [3]. With small values of inter-channel gaps, the influence of the effect of four-wave mixing and cross-phase modulation increases, which begins to limit the maximum range of non-regenerative information transmission due to a decrease in the signal-to-noise ratio. A small inter-channel distance can also limit the ability to transmit information at a high channel rate, since there is an overlap in the spectra of adjacent channels (Fig. 1).

Interchannel interference together with intersymbol interference presents itself as serious influencing factors that degrade the quality of signal reception. They generate two-dimensional (2D) interference that must be efficiently processed by digital signal processing at the receiver [4].

In the absence of interchannel distortion, intersymbol distortion can be compensated for by an adaptive equalizer and an FEC decoder. Some form of joint processing of spectrally overlapping asynchronous WDM channels is required to address the combined effects of this interference [4].

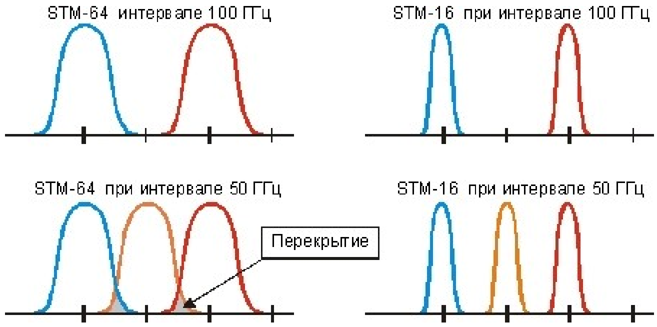

Figure 1. Multiplexing of STM-64 and STM-16 channels at intervals of 100 GHz and 50 GHz

The use of advanced modulation formats and photonic techniques allows signals to be transmitted at a channel rate of 100 Gbit / s in WDM with a fixed grid of 50 GHz. However, for higher traffic rates such as 400 Gbps and 1 Tbps, the required bandwidth using standard modulation formats becomes too wide to fit in a 50 GHz grid. To prevent channel-to-channel crosstalk, one option may be to increase the fixed mesh width from 50 GHz to 100 GHz. The disadvantage of using a wider grid is not only that fewer wavelengths will be transmitted, but also that channels with low speed channels will use a grid up to 100 GHz each, which leads to a decrease in the efficiency of spectrum resource use. Another attempt would be to use higher spectral efficiency (SE) modulation formats such as QPSK and QAM. This option leads to a reduction in the transmission range due to the increased requirements of the optical signal-to-noise ratio (OSNR).

Flexible mesh application

To overcome the aforementioned disadvantages in transition to a transmission rate of more than 100 Gbit / s, Flex Grid technology was proposed, defined in ITU-T G.694.1 recommendation based on 12.5 GHz channel spectral separation (Fig. 2). Thanks to this technique, it becomes possible to flexibly shape the spectrum of any channel, including a superchannel in a certain range of optical frequencies and to scale capacity of the optical network [5].

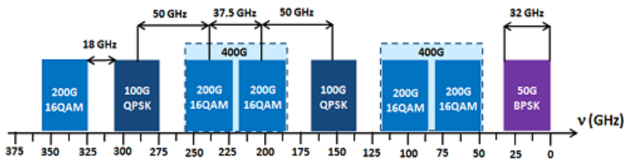

Figure 2. ITU-T G.694.1 flexible frequency grid based on 12.5 GHz grid

The possibility of using Flex Grid technology in commercial systems appeared only after the creation and start of mass production of tunable wavelength selective switches (WSS) using LCoS (liquid crystal on silicon) technology [6].

In [5], a theoretical analysis was carried out comparing flexible mesh and standard mesh, taking into account optical channels between 10 Gbps and 400 Gbps. The characteristics of signals of each type of grid (modulation format, spectrum efficiency in bits / symbol and guard band) are given in Table. 1 derived from data from several transmission studies.

Table 1

Comparison of the use of spectral for different modulation formats

		Fixed Grid 50 GHz		Flexible mesh		
Transmission speed, Gbps	Modulation format	Number of wavelengths	Spectrum, GHz	Number of slot	Spectrum, GHz	Capacity gain
10	NRZ-OOK SE= 1bit / s/Hz	1	50	2	25	100%
40	DP-QPSK SE= 4bit / s/Hz	1	50	2	25	100%
100	DP-QPSK SE= 4bit / s/Hz	1	50	3	37.5	33.3%
400	OFDM-DP-QPSK SE= 4bit / s/Hz	4	200	10	125	60%

As you can see from the table, a 10 Gbps channel using the NRZ-OOK modulation format requires a 25 GHz slot (2 slots) in a flexible mesh, while 50 GHz is used in a fixed mesh case. The remaining channel types (40, 100 and 400 Gbps) use DP-QPSK as the modulation format, reaching a spectral efficiency of 4 bits / symbol. The spectral bandwidth shown in table. 1, takes into account a 7 GHz guard band between optical channels. It also takes into account the effect of forward error correction by increasing the data rate by 12%. In the case of a fixed mesh, the 400 Gbps link requirements are served by four 100 Gbps links, so 4 carriers are used (total 200 GHz in the case of a 50 GHz grid). The values in the last column indicate that using the DP-QPSK modulation format, total FOTS bandwidth can theoretically be increased by 33.3% using 100 Gbps links or 60% using 400 Gbps links.

The overall increase in FOTS throughput using Flex Grid technology, coupled with the ability to transfer speeds in excess of 100 Gbps over long distances, are consecutive reasons for the development of flexible grid technology. However, from the perspective of network operator, it is important to know when this increase in FOTS capacity will represent a viable solution for their networks and how this can be implemented. Depending on the existing fixed grid infrastructure, some solutions may not be supported due to the required spectrum bandwidth. For example, transmissions above 100 Gbps using the DP-QPSK or OFDM-DP-QPSK modulation formats are not possible within the fixed 50 GHz grid. Although there are modulation formats with higher spectral efficiency, such as DP-16-QAM, which can fit into a fixed 50 GHz grid, such formats cause a drop in the maximum transmission range.

Conclusion

As the ICT industry develops, so does the need to increase bandwidth and maintain bandwidth flexibility. Reducing the distance between channels is one of the ways to increase FOTS bandwidth. However, the effect of co-channel interference (the effect of four-wave mixing and cross-phase modulation), leading to a decrease in the signal-to-noise ratio, must be taken into account. Comparison of flexible mesh and fixed network for different bit rates and modulation formats showed the advantage of flexible mesh in bandwidth efficiency. The use of Flex Grid technology, which allows spectrum management, offers the prospect of increasing FOTS operation efficiency at channel speeds above 100 Gbit/s.

References

1. E. Desurvire, C. Kazmierski Science and technology challenges in XXIst century optical communications. *C. R. Physique* 12. 2011. P. 387-416.
2. ITU-T Recommendation G.694.1: Spectral grids for WDM applications: DWDM wavelength grid.
3. E. Connolly et al. Cross-Channel Interference Due to Wavelength Drift of Tunable Lasers in DWDM Networks. *IEEE Photonics Technology Letters*, vol. 19, no. 8, April 15, 2007.
4. F. Lehmann, P. Ramantanis and Y. Frignac "Joint Channel Estimation, Interference Mitigation, and Decoding for WDM Coherent Optical Communications," *IEEE/OSA Journal of Optical Communications and Networking*. Vol. 6. No. 3, pp. 315-325, 2014.
5. V.G. Fokin, R.Z. Ibragimov. Optical systems with terabit and petabit transmission rates. tutorial; Siberian State University of Telecommunications and Informatics; department Multichannel telecommunication and optical systems. Novosibirsk, 2016.162 p.
6. A.V. Leonov, O.E. Naniy, M.A. Sleptsov, V.N. Treshchikov. Trends in the development of long-distance optical systems. *Applied Photonics*. Vol. 3. No.2 (2016).
7. N.N. Slepov. About modern technology WDM and not only. *Foton-Express*. 2007. No. 1. P. 8-16.

COMPARISON OF SEVERAL MODELS FOR CARDIOVASCULAR DISEASES PREDICTION

Wang Yue,
Urumqi University, Urumqi, China, wynfnone@gmail.com

Lilia I. Voronova,
Moscow Technical University of Communications and Informatics, Moscow, Russia, voronova.lilia@ya.ru

Vyacheslav I. Voronov,
Moscow Technical University of Communications and Informatics, Moscow, Russia, vorvi@mail.ru

DOI: 10.36724/2664-066X-2020-6-6-24-28

ABSTRACT

Due to the rapid development of economy, science and technology, the pace of life of people has accelerated and their standard of living has increased. At the same time, the number of various chronic diseases, such as cardiovascular, cerebrovascular and chronic heart diseases, is increasing. These problems seriously affect people's quality of life. Therefore, the problem of predicting cardiovascular diseases has become extremely urgent. The article compares several models for predicting heart disease and evaluates quality of their prognosis.

KEYWORDS: *cardiovascular diseases, random forest, k-nearest neighbors, naive Bayesian classifier, decision tree, artificial neural network, disease prognosis*

INTRODUCTION

According to the World Health Organization, “Cardiovascular disease (CVD) is the leading cause of death worldwide – more people die from CVD every year than from any other disease. In 2030, about 23.6 million people will die from CVD, mainly from heart disease and stroke. These diseases are projected to remain the main single cause of death”[2].

This shows that predicting cardiovascular disease has become extremely important. In the field of disease prediction: S. M. K. Chaitanya et al. [3] proposed the use of artificial neural networks and gravity search algorithms to detect chronic kidney disease. Maryam Tayefi et al. [4] used a decision tree algorithm to create a predictive model of coronary heart disease to identify risk factors associated with coronary heart disease. Ms S. Kalaiarasi et al. [5] proposed using a mobile phone camera to collect facial data from the human body and create an application that can help diagnose and predict skin diseases using the device's image processing and machine learning functions. Pronevska and Claudia [6] use a random forest algorithm to classify biomedical signals, which increases the reliability of medical diagnostics. Patil et al. [7] presented a K-Means cluster algorithm for retrieving data suitable for heart attack from a data warehouse. Resul et al. [8] proposed forecasting using a set of neural networks that combine existing methods to create new models for predicting diseases.

At the Department of Intelligent Systems in Control and Automation of MTUCI, a large scientific and methodological work is being carried out on the use of modern methods of data mining in the educational process within different disciplines, which allows to form the competencies of bachelors and undergraduates in accordance with the challenges of Industria4.0. At the same time, classical approaches are combined and replenished with new methods. Big Data and databases [9], [10], research methods and models for predicting diseases [11], social adaptation of people with disabilities using computer vision [12].

In this paper, we compared several models for predicting cardiovascular disease based on decision trees, random forests, k-nearest neighbors, naive Bayesian classifiers, and artificial neural networks, and evaluated their predictive effects.

I. Data set

The biomedical dataset used in this study was created by Svetlana Ulyanova, Department of Data Science, Ryerson University, Toronto, Canada [13]:

https://www.kaggle.com/sulianova/cardiovascular-disease-dataset

The dataset contains 70,000 data records, including 11 characteristics such as age (age), gender (gender), height (height, cm), weight (weight, kg), systolic blood pressure (ap_hi), diastolic blood pressure (ap_io) , cholesterol (cholesterol, with values of 1: normal; 2: higher than normal; 3: much higher than normal), glucose (gluc, with values of 1: normal; 2: higher than normal; 3: much higher than normal), smoking (smoke), alcohol consumption (alco), and amount of exercise (active). two target outcome classes (cardio), 0: healthy and 1: suffering from cardiovascular disease.

Table 1 shows 21 records from the dataset [13].

Figure 2 shows the correlation between attributes. From the correlation map, it can be seen that cholesterol, blood pressure (ap_hi and ap_low both) and age are closely associated with cardiovascular disease. Glucogen and cholesterol are also closely related. Patients with cardiovascular disease have higher levels of cholesterol and blood sugar, and generally less exercise.

Figure 2. The relationship between every two properties

Table 1

Partial initial data from the set [13]

id	age	gender	height	weight	ap_hi	ap_lo	cholesterol	gluc	smoke	alco	active	cardio
0	18393	2	168	62	110	80	1	1	0	0	1	0
1	20228	1	156	85	140	90	3	1	0	0	1	1
2	18857	1	165	64	130	70	3	1	0	0	0	1
3	17623	2	169	82	150	100	1	1	0	0	1	1
4	17474	1	156	56	100	60	1	1	0	0	0	0
8	21914	1	151	67	120	80	2	2	0	0	0	0
9	22113	1	157	93	130	80	3	1	0	0	1	0
12	22584	2	178	95	130	90	3	3	0	0	1	1
13	17668	1	158	71	110	70	1	1	0	0	1	0
14	19834	1	164	68	110	60	1	1	0	0	0	0
15	22530	1	169	80	120	80	1	1	0	0	1	0
16	18815	2	173	60	120	80	1	1	0	0	1	0
18	14791	2	165	60	120	80	1	1	0	0	0	0
21	19809	1	158	78	110	70	1	1	0	0	1	0
23	14532	2	181	95	130	90	1	1	1	1	1	0
24	16782	2	172	112	120	80	1	1	0	0	0	1
25	21296	1	170	75	130	70	1	1	0	0	0	0
27	16747	1	158	52	110	70	1	3	0	0	1	0
28	17482	1	154	68	100	70	1	1	0	0	0	0
29	21755	2	162	56	120	70	1	1	1	0	1	0
30	19778	2	163	83	120	80	1	1	0	0	1	0

II. Data preprocessing

Before performing data operations, you must cleanse and preprocess the data. Detecting and processing outliers can improve the estimation of forecast accuracy.

To compare the properties (age), (height), (weight), (ap_hi), (ap_lo) on the same scale, they must first be standardized. And let's use min-max normalization.

$$X_{norm} = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \quad (1)$$

where X is the data, Xmin is minimum value of data sample, Xmax is the maximum value of data sample.

We randomly select 80% of data records in dataset as the training set for training model, and remaining 20% of data records as test case for testing model.

III. Designing and comparing multiple predictive models

In this work, we mainly use the Scikit-learn library [14] to complete model building based on decision trees, random forests, k-nearest neighbors, naive Bayesian classifiers, and artificial neural networks. Scikit-learn provides a variety of algorithms for Supervised Learning and Unsupervised Learning through an interface to the Python programming language.

• *Decision tre*

The decision tree adopts a greedy top-down algorithm that classifies the samples by choosing the best classification attribute at each node, and then continues the process until the tree can accurately classify the training samples or all properties have been used.

Listing 1. A snippet of program code

```
from sklearn.tree import DecisionTreeClassifier
dec = DecisionTreeClassifier()
dec.fit(x_train, y_train)       //training
scores["Decision tree"] = dec.score(x_test, y_test)
```

• *Random for st*

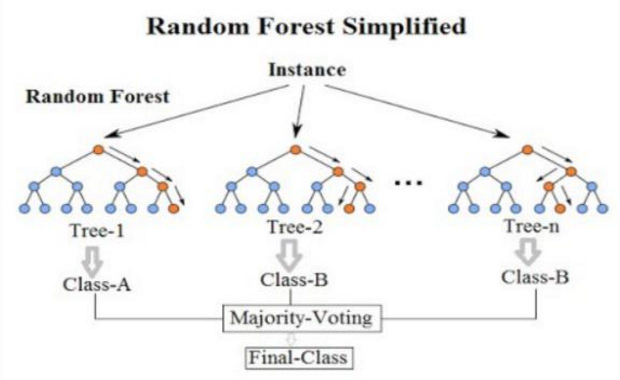

Figure 3. Simplified random forest

The random forest belongs to the bagging algorithm in Ensemble Learning. It is a technique for distinguishing and classifying data using multiple decision trees. Its main unit is a decision tree. He can assess the importance of each variable in categorical data, as well as assess the role of each variable in categorization.

Listing 2. A snippet of program code

```
from sklearn.ensemble import
RandomForestClassifier
ran = RandomForestClassifier(n_estimators=100)
ran .fit(x_train, y_train)    //training
scores["Random forest"] = ran .score(x_test, y_test)
```

• *K-nearest neighbors method*

The k-nearest neighbors method is a nonparametric approach in which the response of a data point is determined by the nature of its k neighbors from the training set. It can be used in both classification settings and regression.

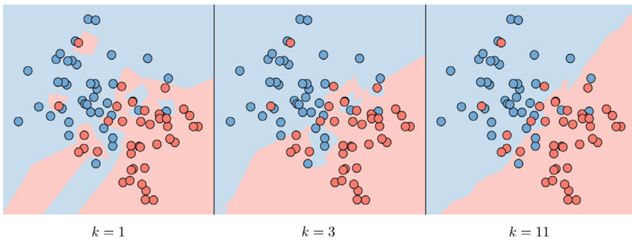

Figure 4. Method of k-nearest neighbors

Listing 3. A snippet of program code

```
from sklearn.neighbors import KNeighborsClassifier
knn = KNeighborsClassifier(n_neighbors=100)
knn .fit(x_train, y_train)    // training
scores["KNN"] = knn .score(x_test, y_test)
```

• *Naive Bayesian Classifier*

Assumption – The Naive Bayesian model assumes that all characteristics of each data point are independent:

$$P(x \mid y) = P(x_1, x_2, \ldots \mid y) = P(x_1 \mid y) \ldots = \prod_{i=1}^{n} P(x_i \mid y) \quad (2)$$

Solutions – The maximum log likelihood gives the following solutions with $k \in \{0,1\}$, $l \in [1, L]$

$$P(x \mid y) = P(y = k) = \frac{1}{m} \times \#\{j \mid y^{(j)} = k\} \quad (3)$$

and

$$P(x \mid y) = P(x_i = l \mid y = k) = \frac{\#\{j \mid y^{(j)} = k \text{ and } x_i^{(j)} = l\}}{\#\{j \mid y^{(j)} = k\}} \quad (4)$$

Listing 4. A snippet of program code

```
from sklearn.naive_bayes import GaussianNB
naive = GaussianNB()
naive.fit(x_train, y_train)      //training
scores["Naive bayes"] = naive.score(x_test, y_test)
```

• Artificial eural network

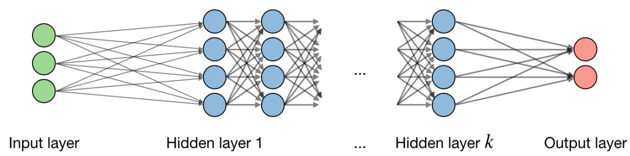

Figure 5. Simplified artificial neural network

Neural networks are a class of models built with layers. Marking the i-th layer of the network and the j-th hidden unit of the level, we get:

$$z_j^{|i|} = w_j^{|i|^T} x + b_j^{|i|} \quad (5)$$

where w-weight, b-bias, z-output respectively.

Examples of the use of neural network modeling in medical problems [16], [17].

Cross entropy loss. In the context of neural networks, cross-entropy loss $L(z, y)$ is usually used, which is defined as follows:

$$L(z,y) = -[y\log(z) + (1-y)\log(1-z)] \qquad (6)$$

Learning rate – the rate of learning, often referred to α, and sometimes η, indicates how fast the weights are updated. This can be corrected or modified adaptively.

We first use the training set separately to train models based on Decision tree, Random forest, k-nearest neighbors (KNN) method, naive bayes classifiers (Naive bayes), and then use the test set to validate these models. The test results are shown in Fig. 6:

	Accuracy Score
Random forest	0.715953
KNN	0.713054
Naive bayes	0.630862
Decision tree	0.630209

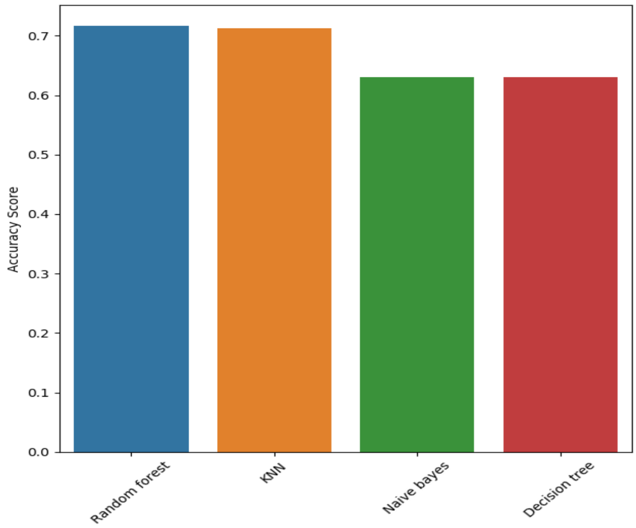

Figure 6. Accuracy test results for multiple models

From the above comparison, it can be concluded that among these models, the random forest model has the highest test accuracy.

Listing 5. A snippet of the program code

```
from sklearn.model_selection import GridSearchCV
grid = {"n_estimators" : np.arange(10,150,10)}
ran_cv = GridSearchCV(ran, grid, cv=3)
ran_cv.fit(x_train,y_train)
print("Tuned hyperparameter n_estimators: {}".format(ran_cv.best_params_))
print("Best score: {}".format(ran_cv.best_score_))
```

Since we used the default parameters in the above model, next we need to determine the optimal value of random forest model parameter using a grid search algorithm to get the best prediction accuracy.

```
Tuned hyperparameter n_estimators: {'n_estimators': 130}
Best score: 0.7159967265417674
```

Next, we evaluate the effect of predicting an artificial neural network model on this dataset. First, we set the serialization model, set the number of neural network layers, select the activation functions sigmod (Logistic), ReLU (Linear Rectifier) and set the learning rate to 0.002.

Listing 6. A snippet of program code

```
model = tf.keras.models.Sequential()
model.add(tf.keras.layers.Dense(6, input_dim=12, activation='relu'))
model.add(tf.keras.layers.Dense(1, activation='sigmoid'))
optimizer = RMSprop(learning_rate=0.002)
model.compile(loss='binary_crossentropy', metrics=['accuracy'], optimizer=optimizer)
model.fit(x=x_train, y=y_train.values,
        batch_size=1024,epochs=1500,
        verbose=0,validation_data=(x_test,y_test.values),
        callbacks=[learning_rate_reduction,es],
        shuffle=True)
```

```
13797/13797 - 0s - loss: 0.5678 - acc: 0.7226
```

The accuracy of testing the trained model reaches 0.7226.Figure 7 shows that the accuracy increases with the number of training iterations. When the number of iterations reaches 1200, the accuracy does not increase significantly. Figure 8 shows that the loss function decreases with the number of training iterations. When the number of iterations reaches 1200, the loss function decreases slightly.

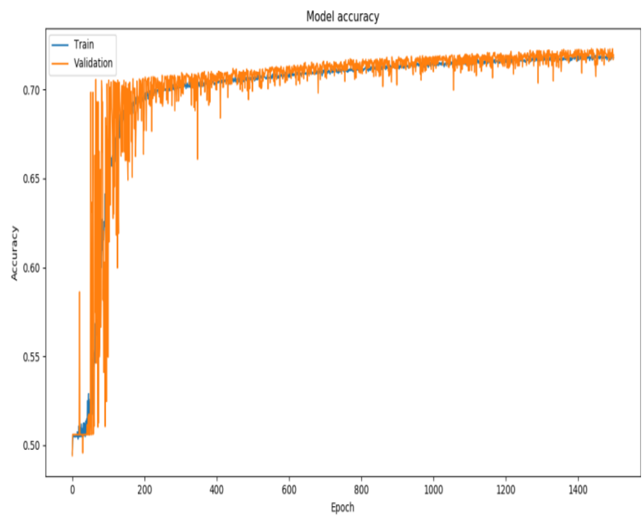

Figure 7. Accuracy changes with the number of training iterations

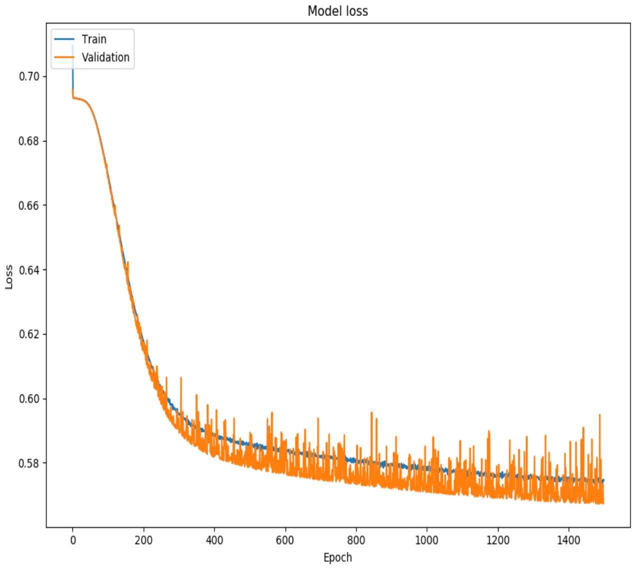

Figure 8. The loss function changes with the number of training iterations

IV. Test results

Table 2

Test result

Model types	Точность
Artificial neural network	0.7226
Random forest	0.7160
K-nearest neighbors method	0.7130
Naive Bayesian Classifier	0.6308
Decision tree	0.6302

Table 2 shows the test results of several models, of which models based on the artificial neural network, random forest, and k-nearest neighbors method have good prediction accuracy, of which artificial neural networks are slightly better than the other two.

Conclusion

This article describes the construction and comparison of several different prediction models applied to a dataset of cardiovascular disease. The forecasting accuracy of the model based on the artificial neural network reached 72%, which was slightly better than in other models. The authors believe that improving the quality of results is advisable to complicate the structure of neural network, improve the training set, and use methods to speed up computations.

References

1. Mortality statistics in Moscow, Moscow City Ritual Service, April 29, 2019 [Electronic resource] https://ritual.ru/poleznaya-informacia/articles/statistika-smertnosti-v-moskve/
2. About cardiovascular diseases, World Health Organization, [Electronic resource] https://www.who.int/cardiovascular_diseases/about_cvd/ru/
3. S. M. K. Chaitanya, P. Rajesh Kumar, Innovations in Electronics and Communication Engineering, January 2019, pp. 441-448.
4. Maryam Tayefi, Mohammad Tajfard, Sara Saffar, et al. Hs-CRP is strongly associated with coronary heart disease(CHD): A data mining approach using decision tree algorithm. *Computer Methods and Programs in Biomedicine*, 2017, 141(4), pp. 105-109.
5. Mrs. S. Kalaiarasi, Harsh Kumar, Sourav Patra,Dermatological Disease Detection using Image Processing and Neural Networks. *International Journal of Computer Science and Mobile Applications*. Vol. 6. Issue. 4, April 2018, pp. 109-118.
6. Proniewska, Klaudia. Data mining with Random Forests as a methodology for biomedical signal classification. Bio-Algorithms and Med-Systems, 2016, 12(2), pp. 89-92.
7. Patil S.B., Kumaraswamy Y.S. Intelligent and effective heart attack prediction system using data mining and artificial neural network[J]. *European Journal of Scientific Research*, 2009, 31(4), pp. 642-656.
8. Resul D., Turkoglu I., Sengur A. Effective diagnosis of heart disease through neural networks ensembles. *International Journal of Expert Systems with Applications*, 2009(36), pp. 7675-7680.
9. Voronova L.I. Teaching aid for the preparation and execution of course projects in the discipline of database technology. Moscow, 2016.
10. Voronov V.I., Voronova L.I., Usachev V.A. Development of a lab workshop for big data processing using hadoop. Moscow, 2018. 49 p.
11. Artemov M.D., Voronova L.I., Voronov V.I., Goncharenko A.A., Yezhov A.A. A software package for sign language recognition based on structural and parametric adaptation of a convolutional neural network. Certificate of registration of the computer program RU 2018666854, 21.12.2018. Application No. 2018664380 dated 13.12.2018.
12. Voronov V., Strelnikov V., Voronova L., Trunov A., Vovik A. Faces 2d-recognition and identification using the hog descriptors method. *Conference of Open Innovations Association, FRUCT.* 2019. No. 24. P. 783-789.
13. Cardiovascular Disease Dataset, https://www.kaggle.com/sulianova/cardiovascular-disease-dataset.
14. Scikit-learn, Machine Learning in Python – [Electronic resource] https://scikit-learn.org/stable/
15. CS 229 - Machine Learning – [Electronic resource] https://stanford.edu/~shervine/teaching/cs-229/cheatsheet-supervised-learning#other.
16. Wang Yue, Voronova L.I., Classification of the type of disease of the spine using a neural network. *Information Society Technologies. Materials of the XIII International Industrial Scientific and Technical Conference*. 2019. P. 407-410.
17. Kesyan G.R., Voronova L.I., Trunov A.S. Predicting the presence of diabetes mellitus using neural networks. *Information Society Technologies. Materials of the XIII International Industrial Scientific and Technical Conference*. 2019. P. 438-440.

DIGEST OF BUSINESS INDEX NORTH. CONNECTIVITY IN THE NORTH (2018-2020)

A periodic report with insight to business activity and opportunities in the Arctic

Erlend Bullveg, (*Chair of the BIN Project Board*), Nord University Business School, Norway,
Erlend.Bullvag@nord.no

Andrey Mineev, (*BIN project coordinator*), High North Center for Business, Nord University Business School, Norway
Andrey.Mineev@nord.no

ABSTRACT

Business Index North (BIN) is a project that contributes to sustainable development and value creation in the Arctic. The overall goal is to set up a recurring, knowledge-based, systematic information tool for stakeholders. This is the second issue of the "Business Index North" analytical report that focuses on the BIN area, including ten northern regions of Norway (Finnmark, Troms, Nordland), Sweden (Norrbotten and Vдsterbotten), Finland (Lapland, Northern Ostrobothnia and Kainuu), and North-West Russia (Murmansk Region and Arkhangelsk Region without the Nenets Autonomous Okrug). For the third issue of the report we would like to include more territories of the Russian High North, as well as Alaska and the Northern territories of Canada. The main implementing partner is the High North Center for Business and Governance at Nord University Business School. Nordland County Council and The Norwegian Ministry of Foreign Affairs provide basic funding for the BIN project.

Chapter Connectivity

– Basic fixed broadband is available to 95% of households in the Nordic BIN regions and to 75% of households in the Russian BIN regions.

– The regions of Troms, Nordland (Norway) and Norrbotten (Sweden) lag behind their country averages in 100 Mbps fixed broadband availability by 8 percentage points and 7 percentage points respectively, while the Finnish regions of Northern Ostrobothnia, Kainuu and Lapland outperform Finland's average by 8 percentage points.

– Mobile broadband coverage (3-4G) is good over all populated places in the BIN area. In terms of territorial coverage in 2016 the BIN regions in Norway had the best coverage lagging behind the national average by only 3 percentage points, Swedish BIN regions lagged behind by 14 percentage points and the Russian BIN regions lagged 21 percentage points behind their respective national averages.

– The BIN area has no direct connection to Europe and North America via subsea data cables. A number of landing points of data cables to Europe are on the coast of South Norway, South Sweden and South Finland. North-West Russia has one subsea data cable to Finland. Direct transAtlantic data traffic between Europe and North America proceeds through 12 submarine data cable systems landing in Denmark, UK, The Netherlands, Germany, France, Spain and Portugal.

This chapter covers the living conditions of people in the BIN area and business infrastructure in terms of access to fixed and mobile broadband. The results reveal common needs for broadband statistic information and make recommendations and highlight implications for policy-makers and investors. This chapter reveals that basic fixed broadband is available to 95% of households in the Nordic BIN regions and to 75% of households in the Russian BIN regions. The BIN regions in Norway and Sweden exhibit higher levels of fixed broadband availability than do the BIN regions in Finland, although the levels in these regions differ with respect to performance in comparison to the respective national levels in each case. Fixed broadband is affordable, with costs ranging from 1.6 to 3% of annual disposable income. When it comes to mobile broadband, 4G coverage with reliable backup of 3G is provided everywhere people live in all the BIN regions. Territory coverage is significantly lower than average for the BIN countries. Further, the BIN area has no direct connection to Europe, North America or Asia via subsea data cables. There are cable project initiatives, including Havfrue, Midgerdsormen and Arctic Connect, which may directly affect the BIN area and play an enabling role in its development.

KEYWORDS: *Connectivity in the North, broadband, subsea data cables, Arctic region.*

Information about BIN Project Board (www.businessindexnorth.com)

Erlend Bullveg, *(Chair of the BIN Project Board)PhD, Dean at Nord University Business School, Norway*
Andrey Mineev, *(BIN project coordinator), PhD, Researcher at the High North Center for Business, Nord University Business School, Norway*

Connectivity is recognized as a prerequisite for economic development in the Arctic. Finland's Chairmanship of the Arctic Council 2017-2019 sets connectivity as one of the priority areas. Access to broadband is essential for connectivity as it serves the needs of business, communities and research.

Telecommunication tower and Aurora Borealis.
Photo: shutterstock

Access to broadband facilitates the development of e-health and digital education. Business opportunities in the Arctic involving shipping, oil and gas, data centres, mining and service industries all need reliable connectivity solutions. The Nordic BIN – countries Norway, Sweden and Finland – have developed broadband plans and digital agendas; however, connectivity in the Arctic regions requires separate attention.

The Arctic Economic Council reports provide an overview of the challenges and ways forward to develop Arctic broadband infrastructure. In this chapter the focus is on the measurable and comparable development of connectivity in the Arctic in terms of the availability, quality and affordability of fixed broadband (including all the main fixed-line broadband access technologies). Furthermore, availability of mobile broadband is reported. Analysis of broadband development projects in the Arctic further highlights drivers and success factors for improving connectivity in the Arctic.

We address connectivity for people and for business in the BIN area. The indicators used in this chapter come from broadband statistics on households. The use of such statistics is well suited for purposes of highlighting people's universal access to basic infrastructure and Internet. The needs of businesses for Internet might vary, but basic household broadband offerings would suffice for small and medium enterprises (SMEs) as the speeds also satisfy the needs of these customers. Broadband speed is usually measured in Mbps (megabits per second), where a high number means faster downloads and uploads when using cloud services, rapid streaming of music or video and smoother video calls. A broadband speed of 100 Mbps would be considered sufficient for SMEs, but larger firms require higher speeds and bandwidths.

Analysis of subsea cable projects explores new potential for increased connectivity for business in the BIN area. This chapter describes living conditions of people in the BIN area in terms of access to fixed and mobile broadband, identifies universal needs for broadband statistic information and presents implications for policy makers and investors.

Indicators used:

Availability of fixed broadband (1) shows the proportion of households with easy Internet access, whether they use it or not. It shows investments in basic infrastructure and people's universal Internet, without measuring actual usage. Quality of fixed broadband is measured in terms of the availability of speeds of 30 Mbps(2) and 100 Mbps. This indicator demonstrates how well the BIN area meets the broadband coverage objectives of the EU Member States: universal broadband coverage with speeds of at least 30 Mbps by 2020 and broadband coverage of 50% of households with speeds of at least 100 Mbps by 2020.

Affordability of fixed broadband is measured by price level and by its percentage of average national income per capita. This indicator shows how well BIN area meets the targets of the UN Broadband Commission, namely that by 2025, entry-level broadband services should be made affordable in developing countries at less than 2% of monthly Gross National Income (GNI) per capita. Availability of mobile broadband demonstrates mobile broadband availability in terms of population and area coverage. Map of potential subsea cable projects illustrates subsea cable initiatives with a potential effect on the BIN area.

1. A ording to Eurostat, broadband refers to telecommunications in which a wide band of frequencies is available to send data. Broadband telecommunication lines or connections are defined as those transporting data at high speeds, with a speed of data transfer for uploading and downloading data (also called capacity) equal to or higher than 144 kbit/s (kilobits per second). In the Russian statistics minimum speed of broadband is 256 kbit/s.
2. Mbps and Mbit/s are used interchangeably.

Findings:

AVAILABILITY

– asic fixed broadband was available to 95% of households in the Nordic BIN regions and in 75% of households in the Russian BIN regions.

QUALITY

– T e target of the EU Digital Agenda for broadband with at least 100 Mbps per second for at least 50% of households by 2020 was already achieved in the Nordic BIN regions in 2016. The target of 30 Mbps for all is yet to be achieved.

– The IN regions in Norway and Sweden exhibit higher levels of quality fixed broadband availability than in Finland

– The regions of Troms, Nordland (Norway) and Norrbotten (Sweden) lag behind their country averages in 100 Mbps fixed broadband availability by 8 percentage points and 7 percentage points respectively, while the Finnish regions of Northern

Ostrobothnia, Kainuu and Lapland outperform Finland's average by 8 percentage points.

AFFORDABILITY

- Fixed broadband is affordable in the BIN area, with broadband expenses constituting from 1.6 to 3% of annual disposable income. Norway has the most expensive broadband, followed by Sweden and Finland. There is no significant price disparity between the Finnish BIN regions and Finland as a whole.

– urmansk Region has more expensive fixed broadband than the Northwestern Federal District in Russia.

MOBILE BROADBAND

– In 2016 the BIN regions in Norway had the best mobile broadband coverage lagging behind the national average by only 3 percentage points, Swedish BIN regions lagged behind by 14 percentage points and the Russian BIN regions lagged 21 percentage points behind their corresponding national averages.

SUBSEA CABLE INITIATIVES

– The BIN region requires improved connectivity with the USA and Asia by subsea fibre cable. Capital-intensive projects demand careful consortium building and secured financing from the initial stage outset. The role of the governments should be considered in securing connectivity in the Arctic BIN area.

3. In the Russian statistics, broadband is defined as access to internet with download speed of at least 256 kbit per second. In the statistics for the Nordic countries, according to Eurostat, the minimum broadband speed is 144 kbit per second.

Availability of fixed broadband in the BIN area Figure 1 shows the country specific shares of households with fixed broadband access. While Norway was the first to reach the 95% threshold in 2016, Sweden and Finland approached it in 2017. Russia is lagging behind by 20 percentage points as of 2016. Figure2 shows that there is no disparity between BIN regions and their country averages (the share for both North-West Russia and the Murmansk and Arkhangelsk region is about 75%; for the Nordic countries and their corresponding BIN regions the shares are close to 95%). The development in North-West Russia in 2016 was at the 2009-10 level of the neighboring Nordic countries and their BIN regions. Today the difference in the share of households with broadband access between the Nordic BIN regions and North-West Russia is about 20 percentage points. In Russia, priority in extending Internet availability of at least 10 Mbps is given to settlements with a population of at least 250 people.

When interpreting the results in Figures 1 and 2 one should remember that availability of fixed broadband meeting the minimum speed requirement is considered (access to internet with download speed at least 256 kbit per second in Russia and with minimum speed is 144 kbit per second in Nordic BIN area). See Table 1 for speed comparisons.

In the Russian statistics, broadband is defined as access to internet with download speed of at least 256 kbit per second. In the statistics for the Nordic countries, according to Eurostat, the minimum broadband speed is 144 kbit per second.

The Digital Agenda presented by the European Commission proposes to better exploit the potential of information and communication technologies (ICTs) in order to foster innovation, economic growth and progress.

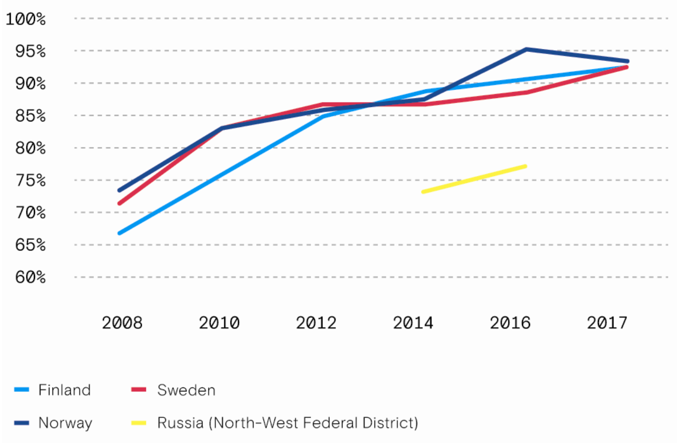

Figure 1. BIN countries-share of households with broadband access, %
Data sources: Eurostat, Rosstat

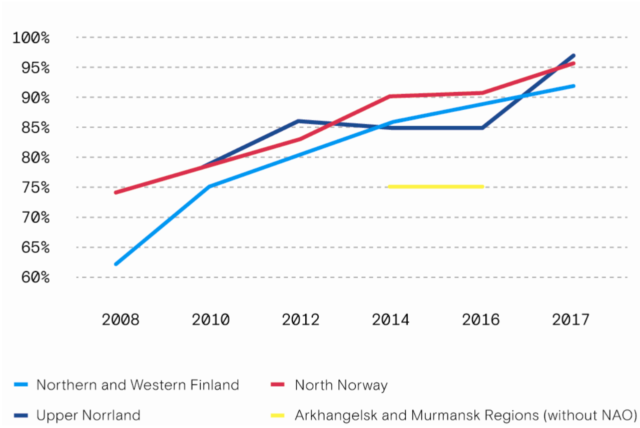

Figure 2. BIN regions-share of households with broadband access, %
Data sources: Eurostat, Rosstat

When it comes to broadband, the Digital Agenda has the following targets:

– All in Europe hall have access to internet with speed over 30Mbps per second by 2020 as the latest.

– 0% of all households in Europe shall have internet subscription with speed more than 100 Mbps by 2020.

In order to give some indication of what these speeds mean for the user, Table 1 compares the broadband speed required for downloading a 5-minnute video and a 2-hour movie over internet by using theoretical calculation. When using an internet connection with 100 Mbps it takes 1.5 min to download a 2-hour movie, while using 256 kbits it would take 9 h and 19 minutes.

Broadband speed comparison

Content	Size	256 kbits	1Mbps	20 Mbps	100 Mbps
5 min video	30 MB	16 min	3 min	13 s	2.5 s
2 h movie	1-1.5 GB	9 h 19 min	2 h	10.5 min	1.5 min

(Source: fastmetrics.com)

Availability of fixed broadband, % of households

Figure 3.1 shows the quality of fixed broadband availability in the BIN area. The target of the EU Digital Agenda for broadband with at least 100 Mb per second for at least 50% of households by 2020 was already achieved in the BIN regions of Nordic countries in 2016. The target of 30 Mbps second for all was yet to be achieved.

Percentages just represent the possibility to acquire broadband (infrastructure in place). The total level of broadband accessibility in Finland is significantly lower than in Sweden and Norway, lagging by 22% for 30 Mbps and 25.5% for 100 Mbps compared to the average for Norway and Sweden. The regions of Västerbotten, Finnmark and Kainuu are among the best performing regions in their respective countries in terms of access to internet with at least 100 Mbps. The differences across countries are explained by country-specific initiatives to support fibre enabled Internet availability. In Finland, the commercial bias has been more toward mobile network development. In Sweden state aid coupled with regional broadband coordinators acting as the link between the regional and municipal level and the market actors deploying broadband infrastructure proved to be efficient in achieving availability of high quality broadband. In Norway there have been more public financial support schemes available in order to cover the costs of the "last mile" of infrastructure in rural areas (5). No comparable statistics are available for Russia.

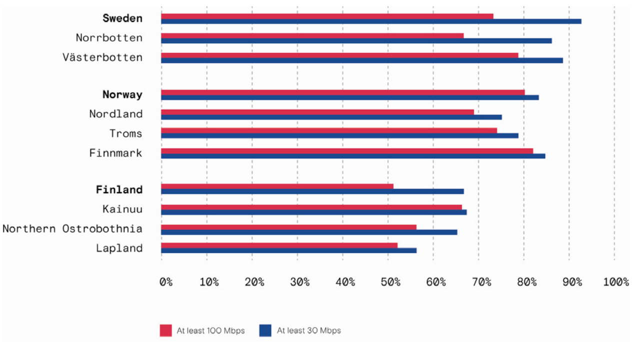

Figure 3.1. Data sources: Finnish Communications Regulatory Authority, Norwegian Communications Authority, Swedish Post and Telecom Authority 5.
ACS Telecoms REPORT

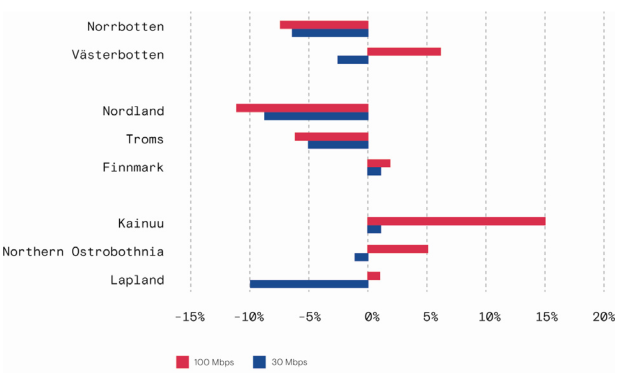

Figure 3.2. Difference in fixed broadband availability by speed compared to country average, %

Figure 3.2 shows that six out of the eight BIN regions underperform in 30 Mbps availability compared to their respective country averages, ranging from a 10 percentage points gap in availability in Lapland to a 1 percentage point gap in Northern Ostrobothnia.

In Sweden, Norrbotten region underperforms in both 30 Mbps and 100 Mbps availability, while Västerbotten outperformed by 6 percentage points in 100 Mbps compared to the Swedish average (see Figure 3). In Norway, both regions of Nordland and Troms underperformed in fixed broadband availability, especially in 100 Mbps speed Nordland lags behind by 11 percentage points and Troms by 6 percentage points. Finnmark region performed slightly better than the Norwegian average. In Finland the BIN regions of Kainuu (15 percentage points), Northern Ostrobothnia (5 percentage points) and Lapland (10 percentage points) outperform Finland's average in 100 Mbps broadband availability equaling 51%, which is considerably lower than for the Swedish and Norwegian BIN regions. There is need to address fixed broadband disparities in the BIN regions.

Subscription price per month, minimum 100 Mbps, lowest price offer in EUR

Figure 4 demonstrates that subscription prices for fixed broadband with at least 100 Mbps second differ widely among the BIN countries. Norway has the highest price and Russia has the lowest. Prices in Sweden and Finland fall in between. Unfortunately, we could not find detailed statistics for the Norwegian and Swedish BIN regions. However, according to the experts, there is no significant difference between regions within the countries. In addition to the subscription price, users often have to pay an opening fee (not shown in the figure). The average level of the opening fee differs greatly between the countries. While in Norway the opening fee is around 564 EUR, in Sweden and in Finland, it is about 1,750 EUR. This means that for first-time users of high speed broadband (100 Mbps is available on fibre lines), the total cost in Norway is lower than in Sweden and Finland if considering a two-year plan. The opening fee in Russia can be up to 500 EUR if there is no fibre cable connected to a house. However, there is a connection to most of blocks, and if there is none people normally do not go for it but opt for a wireless connection plan.

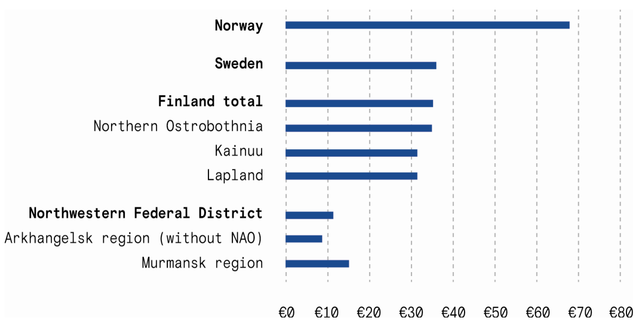

Figure 4.
Data sources: Finnish Communications Regulatory Authority, Swedish Post and Telecom Authority, Internet providers

Affordability of fixed broadband by speed: annual subscription price as % of annual net income

Broadband of 100 mbps is most affordable in Finland, and amounting to 1.9 % of annual net income, followed by Sweden (2%). The most expensive high-speed broadband is in Norway, 3% of annual net income. Affordability of 30 Mbps broadband was under 2% in all BIN regions and their corresponding countries. The greatest affordability disparities are in the Murmansk region.

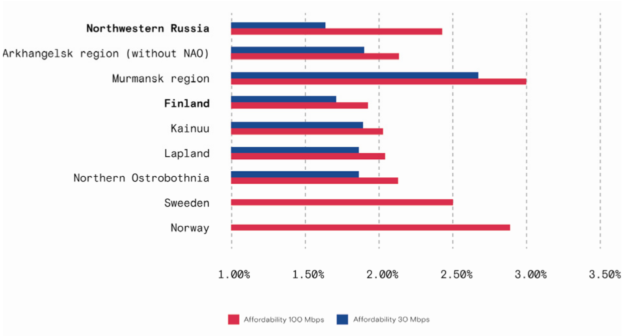

Figure 5. Affordability of fixed broadband as percentage of annual net income

Data sources: Finnish Communications Regulatory Authority, Swedish Post and Telecom Authority, Internet providers, Statistics offices in Norway, Sweden, Finland, Russia.

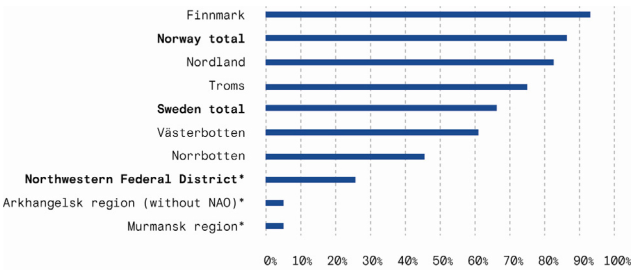

Figure 6. 4G area coverage in the BIN regions, % of own territories *(3-4G coverage for Russian regions)*

Mobile broadband in the BIN area

The estimated share of the population with access to mobile data based on 4G (all mobile networks, outdoor coverage) is close to 100% in all BIN regions in Norway, Sweden and Finland. In the Russian BIN regions of Murmansk, Arkhangelsk and Northwestern District in Russia as a whole most households are covered with 3G and some have 4G. While population coverage with mobile broadband is nearly complete, the area coverage in the BIN regions is much less (Figure 6). The northern territories of Norway, Sweden, Finland are much less covered with 2G, 3G, 4G than the southern parts of these countries.

The situation on the Russian side is even more dramatic – most of the Northwestern Federal District Territory remains uncovered. On the Nordic side, most of the territory covered has 4G and a secure 3G back-up. In Russia most of the territory covered has 2-3G, while 4G is available only in more densely populated places. Mobile networks are developed first in populated areas (see Figure 7). The higher percentage of territory covered – the more dispersed the population in the region is and vice versa. Finnmark in northern Norway has the largest share of own territory covered with 4G – 93%. In general, regions in Norway have a higher share of own territory covered than do Swedish regions. The Russian regions have the lowest share. No comparable Finnish data available.

Area covered with 4G (in 1000s km2) and number of citizens per 1 km2 of this area

Figure 7 shows that Finnmark has the largest area covered with 4G – 45.2 thousand square kilometers. At the same time, the region has the lowest number of people per square kilometer of the area covered with 4G – 2. Figure 7 clearly shows that the BIN regions with the largest 4G covered areas have the lowest number of population per square kilometer of this covered area, and viceversa. Murmansk region has 105 people per square kilometer of the area covered with 3-4G, while this area is only 7. 2 thousand square kilometers. No comparable Finnish data available.

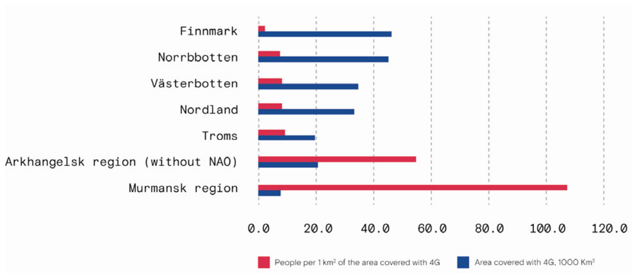

Figure 7.

International subsea fibre initiatives in the Arctic

The needs of modern internet users require fast internet with low latency, meaning short delays in data transmissions. The driving factors behind the need for higher bandwidth are among others increasing cloud driven traffic, IoT developments, Industry 4.0, autonomous vehicles, emergence of 5G technology which offers data transfer up to 150 times faster than the current 4G networks. Subsea fibre cables carry close to 100% of transoceanic voice and data communication.

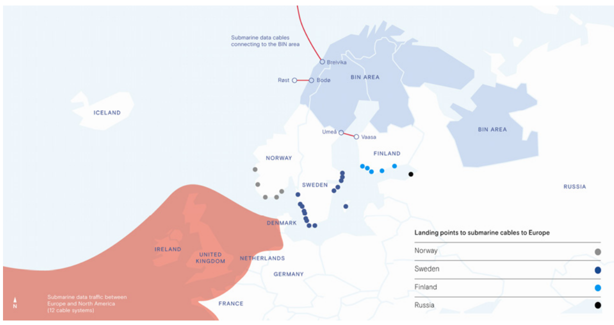

Figure 8.1. The BIN area on the Submarine. Cable Map
Data source: https://www.submarinecablemap.com/
The Submarine Cable Map is a free and regularly updated resource from TeleGeography

Figure 8.1 shows that on the global scale the BIN area has no direct international subsea fibre cables connecting it to the USA or Asia. The lines on the map show the routes of the cables and territories they connect. As of early 2017, there are approximately 428 submarine cables in service around the world. The total number of cables is constantly changing as new cables enter service and older cables are decommissioned. Historically, these cables were built as commercial projects financed by private enterprises rather than governments. The BIN area has no direct subsea cables to the USA or Asia; Direct transatlantic data traffic proceeds through 12 cable systems connecting regions in North America to Denmark, Netherlands, Germany, United Kingdom, Ireland, France, Spain and Portugal. Major subsea cables that connect Finland, Sweden and Norway with the rest of the world have interconnects in continental Europe, which introduces latency into data traffic. All landing points for these cables are in the south of Norway, Sweden, Finland and none in the BIN area; South Sweden has the highest number of landing points for these cables to continental Europe.

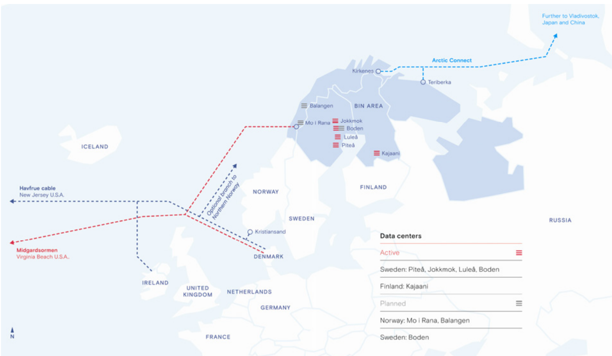

Figure 8.2. International subsea fibre initiatives in the Arctic

The opening of the Arctic sea and operational Northern Sea Route create preconditions for northern subsea cables (see Figure 8.2). The growing business potential of the Arctic requires new subsea cable solutions to improve Arctic connectivity with the rest of the world. In the BIN area, the need for fast connectivity is driven especially by:

– Interest n the BIN area as an attractive place for opening data centres (see datacenters on the map) (*There are datacenter initiatives at initial stage in North-West Russia that are not displayed on the map*) running on low cost green energy, benefitting from cold climate and taxation incentives

– Offshore wind industry

– ining

– O l and gas industries

– ncreased demands of local businesses in cloud computing services

– Tourism and transport industry connectivity needs

– BIN area being a frontrunner in 5G research (*The University of Oulu has a 5G Test Network https://5gtn.fi*)

In our analysis we investigate three BIN area related subsea cable initiatives originating in different countries and at different stages of completion. Attention is paid to how projects are organized and their sources of investment. As a benchmark we investigate two reference projects with potential impact on the BIN area. The main development and success drivers are identified.

BIN area subsea fibre cable projects

Arctic Connect is a cooperative opportunity for PolarNet and Cinia Group Oy to relaunch Polar Net's Russian Optic Trans Arctic Submarine Cable System (Arctic Connect). The Russian Optical Trans-Arctic Submarine Cable System ("R.O.T.A.C.S.") is a Russian-led project that began in the year 2000 and was developed by PolarNet. Midgårdsormen Norwegian-led project seeking to design, build and operate a Norway-centric transatlantic 7,500 kilometre cable system to connect Norway and Sweden to the East Coast of the United States. Specifically, Midgårdsormen proposes to connect Virginia Beach, Virginia to Blaabjerget, Denmark, with a possible connection to Mo i Rana, Norway.

NXTVN´S Oulu Nordic Express Europe proposes a cross-border, Nordic-centric, Gulf of Bothnia bridge connecting cities in the Nordic regions of Finland and Sweden to Norway with onward connections to mainland Europe via submarine and terrestrial networks. NXTVN specializes in Data Center Parks solutions.

Reference subsea fibre cable initiatives

For reference, we use two projects that are in the operational stage and that are likely to affect connectivity in the BIN area. Quintillion brings high-speed Internet access to the North American Arctic through subsea cable. Quintillion is a private operator that contracts to sell capacity on a wholesale basis on its network.

Havfrue subsea cable will run through the North Atlantic connecting mainland Northern Europe to the USA. Optional branch extensions to northern and southern Norway are also included in the design. The first new transatlantic cable in almost two decades.

The analysis served to identify the following success drivers of subsea cable projects:

– Se ured finances

– S rong consortia

– Growing role of OTT players (over-the-top), e.g. Facebook, Google and Amazon as initiators of subsea cable investments.

Recommendations

Overall, connectivity of the BIN area should be addressed at the government level, including the interests of different stakeholders such as communities, businesses and academia.

For Policy

– Providing all hous holds with access to internet with speeds over 30 Mb per second by 2020.

– Improving mobile broadband availability in the BIN area.

– Decreasing discrepancies in broadband affordability in the Russian BIN regions.

– Address ng the needs for increased connectivity by means of subsea cables connecting the BIN area with the USA and Asia.

– Addre sing the needs for increased connectivity using the mix of technologies including satellite solutions.

Next we address the following questions:

• What are the levels of fast bro dband accessibility of 100 Mbps and over in the BIN area?

• What are the levels of ultrafast broadband accessibility of 1 Gbps in the BIN area?

• What technologies are emp oyed in providing fast broadband accessibility?

• Is there any differ nce in broadband accessibility between households and businesses?

• H w is the data centre industry developing in the BIN area?

To address these questions, we collect and analyse comparable data across Finland, Sweden and Norway, but do not include the Russian BIN regions for which such detailed data is not available. Levels of broadband accessibility are compared by using an indicator of fast fixed broadband access of 100 Mbps and more. The EU broadband objectives for 2020 include providing half of European households with connectivity rates of 100 Mbps and by 2025 with access to connectivity offering at least 100 Mbps for all European households.

To assess ultrafast broadband accessibility we collect data on 1 Gbps in the BIN area. Connectivity of 1 Gbps or Gigabit connectivity is needed for educational services online, digitally intensive enterprises, manufacturing systems, ordering and delivery processes, data storage and analytics.

The EU's strategic objective for 2025 is to provide Gigabit connectivity to all main socio-economic drivers such as schools, transport hubs and main providers of public services, as well as digitally intensive enterprises. Additionally, we provide data on broadband accessibility for households and businesses to access how these two groups of users benefit from connectivity in the BIN area. Finally, we collect data to map existing and planned data centre activity in the BIN area.

Key findings:

• Di ital infrastructure in the BIN area is good for supporting the needs of households.

• On aver ge 75% of households in the BIN area have access to high speed broadband of 100Mbs and higher.

• Ultrafast int rnet of 1 Gigabit via fiber optic access is available to 58% of households in the BIN area.

• Sweden leads in providing very highspeed broadband via fiber access to 82% of households in the BIN regions.

• Digital infrastructure acce sibility for business users in Sweden and Norway is on average 10 percentage points lower than for households.

• Data ce tre activity is on the rise in the north.

• Physical digital infrastructure and countries' national support for data centre activity are among the most crucial factors for the future of this industry.

Figure 9 demonstrates that high-speed broadband access is available to an average of 71% of households in the BIN area, which is five percentage points lower than the totals for Sweden, Norway and Finland (76%).

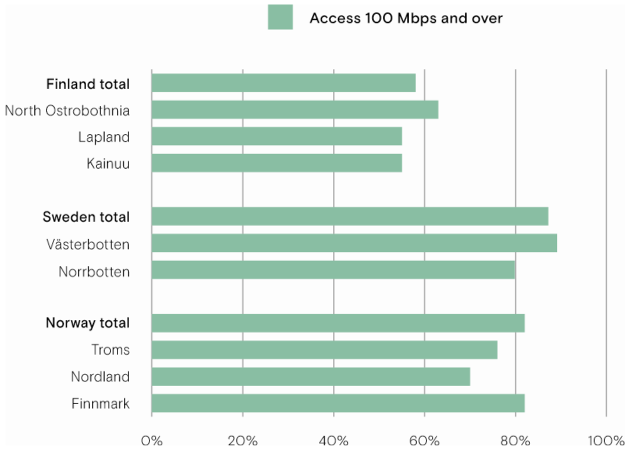

Figure 9. Fixed broadband access 100 Mbps and over, % share of households

Data sources: NKOM, PTS, FICORA (Traficom).

In Norway, Finnmark has the highest fixed broadband access at 82%, with Troms and Nordland reaching respectively 76% and 70%. The topography of the place contributes to the ease of creating infrastructure, reflected in higher accessibility rates in the flat landscape of Finnmark compared to the mountainous terrain in Nordland and Troms.

The low figures for Finland do not necessarily tell the whole story since 89% of households (10% of the area) are covered with at least one mobile network capable of providing 100 Mbps in ideal circumstances. Therefore, in Finland, the lower fixed broadband accessibility is compensated by access to highspeed mobile broadband. All BIN regions have already achieved the EU target of 50% of households with connectivity rates of 100 Mbps by 2020.

Figure 10 demonstrates infrastructural readiness for the provision of very high-speed internet of 1Gbps and more. The data is available for the provision of very high internet access via hybrid fiber-coaxial TV cable (HFC) or via fiber-optic cable (FTTx-fiber to the home).

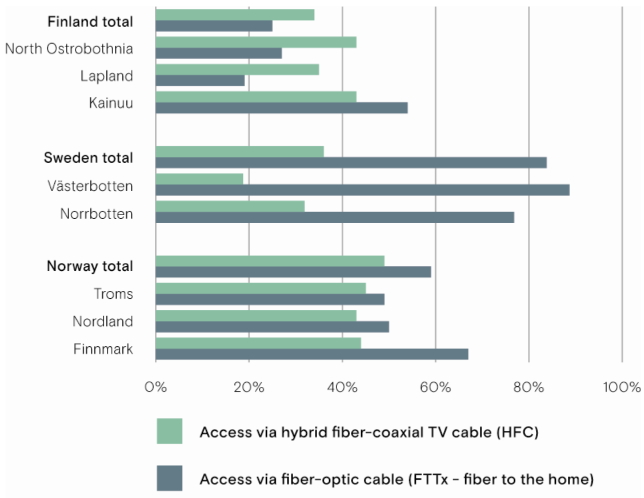

Figure 10. Digital readiness for very high speed broadband 1 Gbps+, by technology, 2017–2018

Data sources: NKOM, PTS, FICORA (Traficom)

The difference between these solutions lies mainly in the upload speed, with fiber-optic cable solutions providing high-speed symmetrical services in both downloading and uploading. Different patterns emerge in the choice of technological solutions across countries, e.g. Sweden has invested heavily in fiber-optic cable (82% of households have accessibility via Access via FTTx-fiber to the home), hence this option leads in readiness to provide very highspeed broadband to households in the BIN area. On average 58% of households in the BIN area have access to fiber-optic cable vs. 56% in the whole of Sweden, Norway and Finland. The high percentage of fiber cable in the BIN area is skewed by the Swedish regions of Norrbotten and Västerbotten with accessibility of 77% and 89% respectively.

Access via hybrid fiber-coaxial TV cable is available to 37% of households in the BIN area vs. 40 % in the whole of Sweden, Norway and Finland. The question of technical solutions is also important to keep in mind when thinking in the long- term; fiber-optic is currently the recommended medium for the link between the core network and the final sub-networks for 5G wireless, so the role of fiber-optical cables is crucial in the uptake of 5G. All BIN regions apart from Finnmark and Västerbotten underperform compared to the country averages in the digital readiness via fiber-optic solutions.

Figure 11 compares fixed broadband accessibility of 100 Mbps and over in households and enterprises in the Swedish and Norwegian BIN regions.

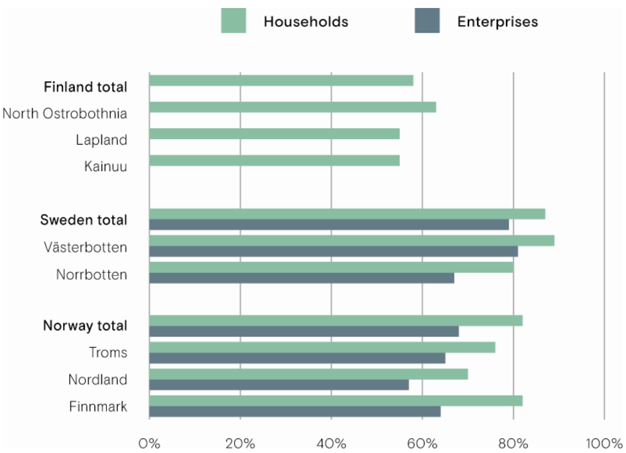

Figure 11. Fixed broadband access by households and enterprises, 100 Mbps and over

The Finnish broadband statistics does not provide data on the enterprise level, but the household statistics serve as a proxy for enterprise broadband accessibility. We observe that on the country level in Norway and Sweden fixed broadband accessibility is on average 12 percentage points lower for enterprises as opposed to households in the same countries, compare 67% vs. 79%. The Norwegian BIN regions have 100 Mbps broadband accessibility for enterprises in the range of 57%-65%.

The difference between 100 Mbps broadband accessibility between households and enterprises on country level as on the BIN regional level can be attributed to the infrastructure installed ready to use within the reach of households, while for enterprises located outside available infrastructure some additional investments may be necessary.

Additionally, smaller enterprises of less than ten employees (i.e. 93% of all enterprises in Finland) can be better represented by using statistics of lower speed internet accessibility.

Data centres in the BIN area

By 2030, the number of devices connected to the internet will have reached 125 billion, up from 27 billion in 20172. These devices rely on the data stored in data centres. The growth in data centres is driven by demand in digital content, mobile computing, Internet of Things (IoT) and cloud computing. Firms worldwide rely more and more on big data and data analysis by external service providers. In 2018 of all Finnish firms using big data 45% relied on external services. In our overview, we examine types of data centres and their capacity. As a measure of capacity we use MW as it best describes power available at data centres.

In our data centre mapping activity in the BIN area we focus on the following types of data centres:

2 ESPAS (2019) Global Trends to 2030: Challenges and Choices for Europe.

3 Definition from Nordic Council of Ministers (2018). Data centre opportunities in the Nordics.

Cloud

Cloud facilities are owned and operated by the cloud companies, which deliver an array of computing services. The larger cloud providers build multisite regional setups with a range of availability zones to ensure low latency and high reliability in the service. The client bases for the cloud companies are all corporate, governmental and individual's applications.

Colocation

A data centre owner sells space, power and cooling to multiple customers in a specific location. The customer typically provides server equipment and the collocation provider hosts it in their data centre by providing space, power and cooling.

Hyperscale

These are large data centre facilities above 20 MW, owned and typically operated by the company they support. They are usually service platforms for social media, search engines, communication & entertainment, artificial intelligence, machine learning and e-commerce5. These data centres are normally located close to the power grid. An example of a hyperscale data centre is the Facebook data centre in Luleå.

Greenfield/brownfield

Greenfield deployment refers to the installation of data centres where previously there was no infrastructure in place, hence necessitating building from scratch. Brownfield development refers to using existing infrastructure not in operation (e.g. old warehouses, factories etc.). Brownfield developments may benefit from existing electricity links. For both types of development brownfield and Greenfield offer specifications of the data centre (e.g.

cloud or hyperscale) determined by the future owner or customer. Table 1 reports data centre development activity in the BIN area by looking at the types of data centres and their operating and planned maximum capacities (where such data is available). All types of data centres are currently being developed in the BIN area, while future sites are mostly greenfield developments. Out of 10 future greenfield development sites, at least four are specifically suitable for hyperscale data centres with electricity capacity expansion of 100 MW up to 300 MW and four planned data centres with expansion capacity above 35 MW.

Capacity data were available for 19 data centres out of 27 data centres, both planned and existing. Operating capacity of existing data centres amounts to 296 MW. If we include operating capacity of planned data centres the result amounts to a total of 635 MW available on short notice. Capacity for expansion data for 13 out of 27 both planned and existing data centres provide an estimate of 1.5 GW operating capacity in the next three years. In terms of energy consumption, it would equal an energy consumption of 13.1 TWh per year. In the chapter of the BIN report devoted to energy, the surplus of the electricity produced in the BIN area amounted to 30 TWh in 2017. Hence, if all planned data centres are built and expanded, the data centre industry may potentially consume up to 44% and more of all energy surplus in the BIN area. Electricity produced from renewable energy sources serves as one of the main attractions for data centre location in the BIN area.

Closeness to the grid and fiber, low cost of electricity, security and political stability and natural cooling conditions are among the selling points encouraging investment in the data centre business in the BIN area. Looking at the regional differences, the Swedish BIN regions are the most prolific in terms of existing and planned data centres compared to their Finnish and Norwegian counterparts.

This is due to the tax incentives introduced by the Swedish government in 2017 to stimulate the ICT sector. As a result newly introduced tax cuts reduced overall electricity prices by around 40 percent for any existing or new data centre greater than 0.5 MW from January 2017 . Additionally, the power companies Vattenfall and Skellefteå Kraft actively promote and facilitate data centre development in the Swedish BIN regions through a Node Pole jointly owned entity established in 2017 serving as a commercial investment and development hub providing dedicated support for investors within the cloud industry and other energy-intensive industries.

Connectivity in the digital age is important for both economic and social development. At the same time, it should not be isolated from the bigger picture, where both positive and negative impacts of increased connectivity are analysed. This type of analysis would intrinsically include environmental impact assessment of both physical infrastructure and future construction projects, e.g. data centres. Additionally, the social component of connectivity should be investigated further, whether better digital infrastructure is the solution to such social problems as loneliness and isolation.

Figure 13. Locations of existing and planned data centres in the BIN area

Figure 13 shows data centre activity in the BIN area in terms of operating and planned data centres. There are currently 11 operational data centres in the BIN area, ranging from Facebook-owned hyperscale data centres in Luleå to the colocation Herman IT data centre in Kajaani. Facebook has confi rmed its additional data centre development in Luleå with at least the same capacity as the previous ones (120 MW) scheduled to be operational in early 2021.

The Norwegian BIN regions of Nordland and Troms have plans for data centre industry expansion in Fauske and Balsfjord, both of which can be used as hyperscale data centres due to the availability of high electricity capacity.

The Swedish BIN regions have seven operational data centres and the data collected reveals that there are 12 potential development sites that can be utilized for data centre industry purposes.

The Finnish regions of North Ostrobothnia and Kainuu collectively have three functioning data centres. Little data is available on the data centre development activity in the Russian BIN regions; in 2018 the administration of Polarnye Zory city signed a co-operation agreement with a Chinese high-tech fi rm for a hyperscale data centre development in the city.

BIN's comment on Covid-19 situation in the Arctic

In this report we assessed the sustainability of the Arctic regions before COVID-19 pandemic hit the world. The spread of the virus and efforts to bring it under control will definitely affect sustainability of the Arctic regions. The scale of the impact will largely depend on the existing conditions for sustainability and governments' responses to the crisis. Although COVID-19 was not the focus of this report, the indicators presented in this report along with previous BIN reports will help readers evaluate vulnerabilities and favourable conditions of the Arctic regions that are now facing pandemic outbreak.

Here we seek to illustrate how indicators can be used to assess vulnerabilities and conditions that may potentially weaken the impact of the virus. The Arctic regions with their low density of population and low urbanization (apart from larger cities in the Russian part of the Arctic) are less exposed to the risk of rapid virus spread. However, there are some places with higher proximity and dense living conditions (i.e. island communities, construction workers settlements) that pose higher infection risks.

Vulnerabilities of the Arctic regions stem from the demographic structure with ageing population and a high proportion of +65-yearolds that are most at risk. Moreover, high proportion of people with chronic diseases and obesity, and mental health issues create additional risks. Historically, the corresponding death rates in the Arctic were already rather high.

In the report we identified negative growth in agricultural and arable land, meaning higher dependency on food produced elsewhere. In the case of supply chain disruptions, this may have negative impact on food security. Tourism in the Arctic is likely to be negatively affected due to fall in demand and imposed travelling restrictions. In particular, hotels, catering, restaurants, entertainment and cultural and creative industries would suffer most from the crisis. Additionally, service providers, retailers are to be potentially negatively impacted. In local communities depending on larger companies, negative impact can be much stronger than in larger cities with distributed economy in the south.

Unemployment in the Arctic regions is expected to increase during the crisis. This will probably strain the Arctic economy. A relative lack of access to capital in the Arctic must be taken into consideration when designing measures for the restart of economic activity.

Broadband access shall be advanced further to meet the demand for remote work and teaching. As a result of pandemic outbreak as well as restrictions imposed by the governments, the Arctic regions are potentially at risk of high unemployment rates, lowering quality of life, depopulation, and less attractive opportunities for investments. On the other hand side, the Arctic regions are so far better off in terms of infection rates. In times of the crisis, we need to build partnerships and learn from each other. Countries have different exit strategies and support mechanisms to re-build the economy. Decisions made as part of the rebuilding plan will have long-lasting effects on all aspects of sustainability.

We therefore challenge authorities to develop a preparedness plan on how to address interconnected risks and achieve sustainability. Evidence from the Arctic regions can be used for targeted measures to build socially, environmentally and economically sustainable Arctic regions during and after the crisis.

Conclusion of the BIN report

We measured and analysed the level of sustainable development in 14 regions in the Arctic Europe including Norway, Sweden, Finland, and Russia. The United Nations Agenda 2030 of sustainable development goals was used as a measurement framework. We used 52 indicators selected from the UN framework under criteria of appropriateness and data availability for the Arctic.

The indicators were grouped into five interlinked pillars of sustainability: People, Society, Economy, Environment and Partnership.

We see big differences between the north and the rest in the four countries of Arctic Europe. Our analysis shows that the situation in the Arctic areas is better only in case of 21% of the indicators. For 34% of the indicators the situation is the same, and about 45% of the indicators describe a situation in the Arctic areas worse than that prevailing in the respective countries as a whole. Specifically, performance is worse on People, Society and Environment indicators. At the same time, Arctic regions in Norway and Sweden are performing better than their respective countries on economic indicators. At aggregate, with the exceptions of the regions of North Ostrobothnia in Finland, and Yamalo-Nenets in Russia, the Arctic areas lag behind their respective countries in terms of sustainable development.

For a more comprehensive view we developed maps and tables where the performance of the Arctic regions can be compared against each other and the corresponding countries.

The Nordic Arctic regions had a total of 29.3 TWh electricity surplus in 2017. There is a need for efficient local use of electricity produced predominantly from renewable sources. The Nordic Arctic region has potential to become attractive for establishing energy-intensive industries.

Business development measured in terms of stock in active enterprises shows growth in the sector of business activities and real estate, and in the hospitality sector, while the number of manufacturing firms is in decline.

The employment growth rate needs to be increased in most of the regions apart from Yamalo-Nenets Autonomous Okrug. The unemployment situation is very different across countries with challenges persisting in Finland and Russia. And now facing Corona, it is expected to reach record levels in all the BIN regions. Job creation, increasing innovative potential and fostering knowledge economy should be on the development agendas of the Arctic regions. Most of the Arctic regions, except North Ostrobothnia, lag behind their countries averages in terms of knowledge infrastructure. There is lack of large companies investing in R&D activities.

Emissions per capita are higher than the respective countries' averages in most of the Arctic Europe regions due to differences in industry structure larger presence of (mining, manufacturing, oil and gas) and climatic conditions. Economic activity conducive to increased emissions needs to be viewed hand-in-hand with wellbeing in the region. It is important to have regionally specific strategies and plans for climate change mitigation that take into consideration all pillars of sustainable development.

Macro-economic indicators stimulating partnership: GDP per capita is lower than the respective national averages for most of the Nordic Arctic regions, but growth rate is higher. For Russia there are big differences between regions in terms of GDP per capita. Regions relying more on natural resources have higher GDP per capita. Given the high inequality of incomes this is a trend limiting partnerships. High level and growth rate of GDP in the regions is associated with overconsumption at the macro-level, which in turn presents problems for environment. Achieving partnerships through macroeconomic stability shall be done in conjunction to human development, sustainable consumption and environmental sustainability.

What brings value to development of the North?

The BIN report provides a comprehensive analysis of sustainable business development in the European north. The report is based on statistical data from multiple sources, using scientific methods, and provides factual and comparable indicators across a set of topics and geographic regions. Several implications and recommendations are presented at the end of each chapter.

We emphasize the value of the people who live in or deal with the north, their livelihoods, and the importance of quality education and job creation. At the same time, successful business activities and economic development are another vital component of value creation.

Thus, value creation involves activities beneficial to both persons and legal entities. In this regard, Business Index North seeks to trace both societal and economic developments in the Arctic and offers a nuanced considered view of how these evolve in combination.

Sometimes we observe success stories associated with positive trends for both, sometimes we face worrying contradictions.

The BIN report is a key tool to understand demographic and human development trends, business activities and opportunities, as well as core conditions such as connectivity, knowledge infrastructure and electricity production in the BIN area. By bringing the pieces of the puzzle together our report offers an intellectual insight into the process of value creation for business and society in the North.

What do we know?

Our report highlights worrying demographic trends whereby population growth in the BIN regions is well below national averages and even negative in the Russian BIN regions. The BIN area fails to attract young adults and families; hence the young population is declining.

Men in the regions are decidedly less well educated than women. Business shows positive trends, especially measured in terms of growth in turnover. Businesses rely on natural resources extracted from the region, but new opportunities in value-added business such as information and communications are on the rise.

Is business development in the BIN area socially sustainable? In the meantime, high economic growth in the BIN area is associated with negative job creation rates. Each year, every 100 new workplaces created in the BIN countries (Norway, Sweden, Finalnd, North West Federal District of Russia considered together) are associated with 15 workplaces lost in the BIN regions of these countries considered together.

The BIN area generates electricity from predominantly green energy sources that create attractive business opportunities for energy-intensive industries such as data centres. Broadband accessibility is good for the needs of households but requires further development for the needs of businesses. Research and development in the business sector is below national averages everywhere except North Ostrobothnia.

Thus, being advangeiously positioned in terms of energy and connectivity, businesses in the BIN area have limited capacity to contribute to a knowledge economy based on intellectuial resources, information and knowhow.

What is next?

Interest in the Arctic and High North among globally operating businesses is driven by the demand for natural resources and the opening up of sea transportation routes. Knowledge about the past gives us some indication of what lies ahead. We expect the growth in economic activity in the north to continue. However, this growth needs to be supportive of the local people and local communities. Any major disruptions, such as the opening of new mines or manufacturing plants, need to be taken into consideration in terms of socio-economic impacts and demographic challenges to the region.

We also strive to show the image of the North not only as a source of extractible resources but as a place full of ideas and unique solutions, which we highlight in a separate BIN Innovations report. The BIN report can serve as a platform for creating a vision for sustainable business in the north that is respectful of nature, people and northern values.

The BIN report is produced in order to serve as an analytical tool for various stakeholders, including decision-makers, media, NGOs, academia and others. The report provides an impartial and independent analysis that can be used for benchmarking and as input for policy-making. It can be used for teaching and raising awareness of the BIN area to the wider world.

We continuously upload new reports to the web site www.businessindexnorth.com. Moreover, we develop data visualization tools and infrastructure maps which can be applied by users according to their specific interests. We believe that all people who would like to contribute to sustainable development of the Arctic will find BIN a practical and insightful tool.

ITU COMPLETES EVALUATION FOR GLOBAL AFFIRMATION OF IMT-2020 TECHNOLOGIES

5G WILL BE THE BACKBONE OF THE GLOBAL DIGITAL ECONOMY

Geneva, 26 November 2020. **The radio interfaces that will be used in full-scale commercial deployment of fifth generation mobile communication networks have achieved global validation with the successful evaluation by ITU of three new technologies that conform with the International Mobile Telecommunications 2020 (IMT-2020) vision and stringent performance requirement.**

The technologies are: 3GPP 5G-SRIT and 3GPP 5G-RIT submitted by the Third Generation Partnership Project (3GPP), and 5Gi submitted by Telecommunications Standards Development Society India (TSDSI).

During the multi-year development and evaluation process by the ITU Radiocommunication Sector (ITU-R), these technologies were deemed to be sufficiently detailed to enable worldwide compatibility of operation and equipment, including roaming. The outcome of this first release of IMT-2020 supporting 5G is a set of terrestrial radio interface specifications which are incorporated into a global standard in the ITU-R Recommendation titled 'Detailed specifications of the radio interfaces of IMT-2020.' This is in final approval to the 193 Member States of ITU.

"IMT-2020 specifications for the fifth generation of mobile communications (5G) will be the backbone of tomorrow's digital economy, transforming lives and leading industry and society into the automated and intelligent world," said Houlin Zhao, ITU Secretary-General. "5G will enable much faster data speeds, reliable connectivity and low latency to international mobile telecommunications (IMT) — all needed for our new global communications ecosystem of connected devices sending vast amounts of data via ultrafast broadband."

Mario Maniewicz, Director of the ITU Radiocommunication Bureau, said: "The successful completion of the evaluation process and the release of this global standard is a significant milestone for the global telecommunication industry and its users. 5G technologies will further enrich the worldwide communications ecosystem, expand the range of innovative applications and support the burgeoning Internet of Things, including machine-to-machine communication." The evaluation of the candidate technologies was not carried out by ITU-R alone. It was a highly collaborative process with substantial input from and coordination with ITU Member States, equipment manufacturers, network operators, and involved national, regional, and international standards development organizations, partnerships, the academic community and fora, since ITU-R provides a unique global framework to discuss the capabilities of new radio technologies.

In early 2012, ITU initiated the development of "IMT for 2020 and beyond", setting the stage for 5G research activities and in 2015 established the vision and requirements for the globalization of 5G. Under ITU's ongoing IMT programme, ITU membership is continuing its long-standing contribution to mobile communications, facilitating its mission to be "committed to connecting the world."?