CTRQ 2015

The Eighth International Conference on Communication Theory, Reliability, and Quality of Service

ISBN: 978-1-61208-402-2

April 19 - 24, 2015

Barcelona, Spain

CTRQ 2015 Editors

Eugen Borcoci, Politehnica University of Bucharest, Romania

CTRQ 2015

Foreword

The Eighth International Conference on Communication Theory, Reliability, and Quality of Service (CTRQ 2015), held between April 19th-24th, 2015 in Barcelona, Spain, continued a series of events focusing on the achievements on communication theory with respect to reliability and quality of service. The conference also brought onto the stage the most recent results in theory and practice on improving network and system reliability, as well as new mechanisms related to quality of service tuned to user profiles.

The processing and transmission speed and increasing memory capacity might be a satisfactory solution on the resources needed to deliver ubiquitous services, under guaranteed reliability and satisfying the desired quality of service. Successful deployment of communication mechanisms guarantees a decent network stability and offers a reasonable control on the quality of service expected by the end users. Recent advances on communication speed, hybrid wired/wireless, network resiliency, delay-tolerant networks and protocols, signal processing and so forth asked for revisiting some aspects of the fundamentals in communication theory. Mainly network and system reliability and quality of service are those that affect the maintenance procedures, on the one hand, and the user satisfaction on service delivery, on the other hand. Reliability assurance and guaranteed quality of services require particular mechanisms that deal with dynamics of system and network changes, as well as with changes in user profiles. The advent of content distribution, IPTV, video-on-demand and other similar services accelerate the demand for reliability and quality of service.

We take here the opportunity to warmly thank all the members of the CTRQ 2015 Technical Program Committee, as well as the numerous reviewers. The creation of such a high quality conference program would not have been possible without their involvement. We also kindly thank all the authors who dedicated much of their time and efforts to contribute to CTRQ 2015. We truly believe that, thanks to all these efforts, the final conference program consisted of top quality contributions.

Also, this event could not have been a reality without the support of many individuals, organizations, and sponsors. We are grateful to the members of the CTRQ 2015 organizing committee for their help in handling the logistics and for their work to make this professional meeting a success.

We hope that CTRQ 2015 was a successful international forum for the exchange of ideas and results between academia and industry and for the promotion of progress in the field of communication theory, reliability and quality of service.

We also hope Barcelona provided a pleasant environment during the conference and everyone saved some time for exploring this beautiful city.

CTRQ 2015 Advisory Committee:

Eugen Borcoci, Politehnica University of Bucharest, Romania
Joel Rodrigues, Instituto de Telecomunicações / University of Beira Interior, Portugal
Pascal Lorenz, University of Haute Alsace, France
Raj Jain, Washington University in St. Louis, USA

CTRQ 2015

Committee

CTRQ Advisory Committee

Eugen Borcoci, Politehnica University of Bucharest, Romania
Joel Rodrigues, Instituto de Telecomunicações / University of Beira Interior, Portugal
Pascal Lorenz, University of Haute Alsace, France
Raj Jain, Washington University in St. Louis, USA

CTRQ 2015 Technical Program Committee

Mandicou Ba, Université de Reims Champagne-Ardenne, France
Dirk Bade, University of Hamburg, Germany
Jasmina Barakovic Husic, BH Telecom – Sarajevo, Bosnia and Herzegovina
Monique Becker, Telecom SudParis - Institut Mines-Telecom, France
Roberto Beraldi, "La Sapienza" University of Rome, Italy
Eugen Borcoci, Politehnica University of Bucharest, Romania
Damian Bulira, Wroclaw University of Technology, Poland
Bruno Checcucci, Perugia University, Italy
Laurent Ciavaglia, Alcatel-lucent, Italy
Mario De Felice, Sapienza University of Rome, Italy
Javier Del Ser Lorente, TECNALIA Research & Innovation, Spain
Michel Diaz, LAAS, France
Manfred Droste, Universität Leipzig, Germany
Ali El Masri, Troyes University of Technology -Troyes, France
Andras Farago, The University of Texas at Dallas - Richardson, USA
Gianluigi Ferrari, University of Parma, Italy
Alexandre Fonte, Polytechnic Institute of Castelo Branco, Portugal & Centre for Informatics and Systems of the University of Coimbra (CISUC) , Portugal
Tulsi Pawan Fowdur, University of Mauritius, Mauritius
Julio César García Alvarez, Universidad Nacional de Colombia Sede Manizales, Colombia
Bogdan Ghita, University of Plymouth, UK
Marc Gilg, Université de Haute Alsace, France
Rita Girão Silva, DEEC-FCTUC / INESC-Coimbra, Portugal
Antti Hakkala, University of Turku, Finland
Bjarne J. Helvik, NTNU, Norway
Robert Ching-Hsien Hsu, Chung Hua University, Taiwan
Ilias Iliadis, IBM Research - Zurich, Switzerland
Mohsen Jahanshahi, Islamic Azad University - Central Tehran Branch, Iran
Brigitte Jaumard, Concordia University, Canada
Sokratis K. Katsikas, University of Piraeus, Greece
Al-Sakib Khan Pathan, International Islamic University Malaysia (IIUM), Malaysia
Wojciech Kmiecik, Wroclaw University of Technology, Poland

Nataša Živic, University of Siegen, Germany
Sladjana Zoric, Deutsche Telekom AG, Bonn, Germany
André Zúquete, University of Aveiro, Portugal

Table of Contents

Seamless Mobility – Improved QoE in Mobile Telephony Through QoS-Sensitive (HOCIS) Convenience Messages

Thomas Schildhauer, Felix Urban, Claudia Nichterlein, Annalies Beck, Daniela Doischer, Ivan Ivanovic
Institute of Electronic Business e.V.
An-Institut der Universität der Künste Berlin (UdK Berlin)
Berlin, Germany
e-mail: felix.urban@ieb.net

***Abstract* — Business communication should become more effective through Unified Communications (UC) solutions. With UC, communication services such as telephony, video conferencing, presence, chat, file sharing, etc. become available independent of location or end device. Seamless mobility - the possibility of remaining connected on a mobile device while the connection is being handed off between different networks - is the most important technical factor in efficient mobile communication, especially in the field of UC telephony, which uses both conventional mobile and Internet Protocol (IP) networks for voice calls. Unforeseeable external influences on the quality of service (QoS) and unforeseeable operating conditions of the mobile technology frequently result in connection disruptions. This wastes the user's time and energy and results in a poor performance rating. Based on the example of a mobile UC solution, this paper presents the requirements placed on mobile UC telephony. The QoS barriers, user expectations, and principles for the design of QoS-sensitive interaction are summarized. An example is provided of how the UC-telephony user's quality of experience (QoE) can be increased on the basis of a Handover Convenience Information Support (HOCIS) by means of implementing QoS-sensitive user interaction.**

Keywords — Unified Communications; Fixed Mobile Convergence (FMC); Handover Convenience Information Service (HOCIS); Quality of Service (QoS); Quality of Experience (QoE); QoS/QoE relationship; Voice over Internet Protocol (VoIP);

I. INTRODUCTION

To achieve reliable seamless mobility, i.e., the possibility for the user to remain continuously connected across various network standards, current and future technologies for the development of a Unified Communications (UC) environment must be able to handle differing – both known and unknown – connection standards, end devices, and application combinations [1]. To ensure this, both network technologies and user navigation must continuously be expanded. The key points of this paper summarize the results of a user-centric design approach that foresees both an expansion of network compatibility and of user experience. It is the researchers' goal to perform an innovation process in developing a user-friendly interface for mobile UC, by the means of a patent-based QoS-sensitive backend technology. The process aims at improving the Quality of Experience (QoE) of mobile workers [2] potentially using UC telephony. Hence, the following Sections summarize the implementation process (analysis, planning, design and development, evaluation, optimization, and implementation) of the so-called "Handover Convenience Information System" patent (HOCIS) [3] in the mobile clients of an existing Voice over Internet Protocol (VoIP) UC environment. The results stated in this paper are excerpts from the transfer research between the TELES FMC+C Innovations GmbH and the Institute of Electronic Business (IEB). To that end, the IEB performed, among other things, primary and secondary research in the form of expert and development workshops, as well as user surveys, focus groups, and analyses of literature. The paper is structured in two main parts. Whereas Sections II to IV provide an overview of current service and design challenges as well as the researchers' findings for a QoS-sensitive user experience design (UXD), Sections V and VI present the case study report of the research-based innovation process, which led from a network technology patent to a user centric service implementation.

II. QOS SERVICE BARRIERS AND RAMIFICATIONS FOR THE USE OF MOBILE UC

UC technology will determine the future of telecommunications [4]. However, according to estimates, there are few concrete figures on the actual use and prevalence of UC technologies. In 2011, Ralf Lehmann spoke of a limited UC prevalence (12 percent) and usage primarily by large corporations [5]. Just as little information is available about the development as on the actual prevalence; there are hardly any figures for the providers' network failure and disruption rates. If it can be assumed that higher data rates and better Wi-Fi connections can be found in offices than the comparatively high fluctuations of mobile networks, the UC technologies in the end depend on the provider's quality of service. External disruptions occur frequently [6]. This holds true especially for mobile data transmission, as here the utilization of the faster Fourth Generation (4G)/Long-term Evolution (LTE) technology is relatively rare. In Germany, for instance, significantly less than 10 percent of mobile communication customers use the new LTE technology, while classical telephony and

 ISBN: 978-1-61208-402-2

Universal Mobile Telecommunications System (UMTS) Internet are much more important for the majority of users [6]. Beyond this differing prevalence of technical standards, the technical challenges, for example, circuit switched fallback (CSFB), must also be overcome [8]. Consequently, unknown disruptions and interruptions, varying connection and network standards and unforeseeable usages must be taken for granted in technology that uses VoIP.

For the user, such connection and bandwidth problems in mobile or stationary VoIP usage have a negative impact or result in a communication breakdown. In such cases, users frequently resort to the simple call with the generally better-established Second Generation (2G)/Global System for Mobile Communications (GSM). For the communication process, this generally results in a change of application and telephone number. Such connection problems waste the user's time and are frustrating. On the other hand, for providers of UC solutions, such situations represent the hazard that the negative perception will be directed at the application. Qualitative interviews by the IEB have shown that in cases of a disruption, a user can rarely distinguish between QoS, bandwidth problems, or software failures.

III. QoS and QoE – Requirements for the Design of QoS-sensitive User Interaction

UC telephony aims to provide increased communication effectiveness, in particular for business telephony. The IEB conducted a user group analysis and surveys of a group of demographically diverse volunteers for the purpose of determining user needs in UC telephony clients. The results showed that the user group of this technology is heterogeneous and their specific requirements can hardly be predicted. This technical and user-specific variability for the expansion of seamless mobility in the professional UC environment shows that mobile telephony applications will have to be able to react flexibly to usage-specific and technical influences. The expert workshops and focus groups demonstrated that along with the requirements for the flexible use of the application, the user also requires a simplification of the technical processes and functional possibilities. Surprisingly, fundamental feedback about telephony status was expressed as one of the primary wishes of the users. It is precisely this example that shows that the service blueprints of conventional mobile telephony apparently contain gaps between the background processes and users' activities. This means that background processes, such as network status and quality or network handovers, have not yet been sufficiently expanded as touch-points (i.e. interaction moments between device and user). The volunteers requested a touch-point for the immediate information about the call being broken off as a service function. Furthermore, those surveyed displayed a great discrepancy in their understanding of technology and in their willingness to learn, as well as in the telephony and cost behavior. For these differing uses and contexts, there was the requirement that universal, comprehensible, but individually configurable service functions would have to be designed. The application-specific user requirements therefore must be implemented with the primary requirement of user friendliness (according to Richter these are: Effectiveness, efficiency, performance, promotion of learning, learnability, potential to be remembered, flexibility, degree of freedom, adjustment, satisfaction, acceptance, task suitability, customization to individual needs, transparency, openness, self-descriptiveness, support, compatibility, conformance to expectations, consistency, uniformity, tolerance, fault tolerance, feedback, manageability, controllability) [9] – and always pursuing the aim of allowing for an interruption-free call.

IV. QoS and QoE – Principles for the Design of QoS-sensitive User Interaction

QoS-sensitive interaction is a unique segment of the field of interaction design that needed be achieved. Whereas video telephony software like Microsoft's Skype offers a visual and acoustic feedback when a call is temporarily interrupted, voice telephony applications rarely give a prompt feedback. The herein described approach on QoS-sensitive user interaction with (HOCIS) convenience messages requires more elements than hold or waiting music. Whilst other UC applications such as Unify's Soft Client, Cisco's Jabber or ComdaSys' MC Client partly provide service functions like network handover between different mobile data networks and handover functions between devices, this approach aims at providing an integrated set of QoS-sensitive support functions for all UC telephony options. This development focuses on a user interaction principles which combines status information, options evaluation and user assistance. From the accompanying research and development, the IEB has determined the following principles that must be pursued in the design of the user interaction and call control: The graphical user interface (GUI) should be based as closely as possible on the specific interaction habits using Android or Apple's mobile Operating System (iOS). Here, however, a distinction between the phones' basic telephony application and the UC application is still required. The handling of the interaction scenarios, during and prior to the call, requires adaptability to individual user preferences. For the sound design of the user interaction, the evaluation with the focus groups showed that a design similar to the telephone sounds the users have already learned, so-called auditory icons (use of such sounds, which are taken from daily life and that are closely linked to the event used for because of previous experiences (for instance: crumple up a paper). A multidimensional conveyance of information is possible, for instance, through indications regarding, size, age, or size of a problem) [10], does not provide enough distinction. For innovative functional parameters like warnings, abstract sounds, i.e., earcons (abstract, synthetic sound events which can create an own, simple language, with variations (density, timbre, level, speed, frequency) of melodic or rhythmic motifs) [11], have proven to be more effective. The design of the sounds takes on a more important role than the conventional design of internal, functional sounds, such as waiting music and call waiting. The fact that most service functions occur during the call means that the

 ISBN: 978-1-61208-402-2

challenge is to provide a sensitive sound design that provides the user with the necessary information about the function and purpose [12] while simultaneously creating a minimum level of distraction. The design of the functional sounds is subject to special volume (amplitude), pitch (frequency), and timbre (spectrum) [13] requirements. They should be audible within complex user contexts, but also distinguishable from the device's own telephony application sounds. These goals can be reached by integrating various functional parameters in a "family" of sounds. The specific requirements for QoS-sensitive interaction design are part of the fundamental need to communicate complex, technical processes in multisensory messages and implement them in options. To properly service user expectations and user demands, the messages must be customizable and adjustable in both form and function. These principles are also applied in the following case study.

V. CASE STUDY: IMPLEMENTATION OF CONVENIENCE MESSAGES IN TELES FMC+C'S UC1000

The requirements for the QoS user interaction described above are the result of recommendations for the implementation of the HOCIS patent in the existing mobile UC clients of the TELES FMC+C company, which is currently available for iOS and Android devices [14]. Key Features of this client are: Single-number service; Seamless call handover between GSM and Wi-Fi, UMTS and LTE with convenience messages; Mobile VoIP, callback and callthrough; Visual voicemail, fax, presence and instant messaging; Extension dialing; Integrated corporate address book; Call transfer from desk phone to mobile phone and vice versa; Mobile access to desk phone features; Customizable to any IP Centrex or IP Centrex. Beyond the functions typical for UC, such as single-number service, instant messaging, presence, integrated company phonebook, etc. the application provides special features, including seamless handover between GSM, Wireless Fidelity (Wi-Fi) technology, UMTS, and LTE connections, and "convenience messages" which communicate information about the status of the telephony service to the user based on a QoS-sensitive evaluation of the telephone background process. The process of evaluating QoS input and translating it into so-called events is handled by the patented HOCIS solution. In particular, QoS information such as signal strength, packet loss, jitter, latency, as well as information about roaming status and network availability or loss, are evaluated, interpreted and reported in events (Table I).

TABLE I. QOS SENSITIVE EVENTS

QoS sensitive events	
#	***Description***
1	Wi-Fi/3G/4G available
2	Wi-Fi/3G/4G with special attributes
3	Slow quality decrease
4	Constant bad quality
5	Disconnect
6	Handover complete
7	Roaming
8	Manual handover

These eight events of the HOCIS patent are the inputs that are available for the QoE/QoS service functions and option menus for the UC1000 mobile clients. These events form the interface between the background processes and the user. The corresponding design of the touch-points provides the user with 10 use cases so as to allow more effective and reliable VoIP-UC telephony. To that end, service functions in the use cases (Table II) distinguish between the network and the device prior to and during the call and the handover. In accordance with a successful or failed service process, the interaction is designed so that continuous communication is maintained.

The aim of QoS-sensitive convenience messages is to offer the user the greatest number of alternative options for an ongoing communication process. The decision as per which events and their corresponding messages were implemented in the frontend of the client and which ones were considered as background processes, in other words as automatic service functions, was based on the researcher's user experience design (UXD) process which included qualitative and quantitative methods. This process consisted of the above-mentioned expert workshops, focus groups and usability tests.

Based on the results of this process, the TELES system of "Convenience Messages" consists of multisensory elements that allow the user to receive the corresponding system messages for any system, always with the aim of providing a more comfortable and reliable telephoning experience. Among the elements included in multisensory convenience messaging (see example in Table III) are: graphic elements, text elements (on-display) sound elements, vibration elements, light-emitting diodes (LED), voice elements.

 ISBN: 978-1-61208-402-2

TABLE II. USE CASES FOR CONVENIENCE MESSAGES

#	**Use cases for convenience messages**
1.1	Cost savings – during call *(successful)*
1.2	Cost savings – during call *(failed)*
1.3	Cost savings – before call *(e.)*
1.4	Cost savings – before call *(f.)*
2.1	Optimization of quality – during call *(e.)*
2.2	Optimization of quality – during call *(f.)*
3.1	Cost savings / Roaming – during call
3.2	Cost savings / Roaming – during call *(f.)*
3.3	Cost savings / Roaming – before call *(e.)*
3.4	Cost savings / Roaming – before call *(f.)*
4.1.1	Forced handover: Change of network – during call *(e.)*
4.1.2	Forced handover: Change of network – during call *(f.)*
4.1.3	Forced handover: Change of network – before call *(e.)*
4.1.4	Forced handover: Change of network – before call *(f.)*
4.2.1	Forced handover: Change of device – during call *(e.)*
4.2.2	Forced handover: Change of device – during call
4.2.3	Forced handover: Change of device – before call *(e.)*
4.2.4	Forced handover: Change of device – before call *(f.)*
5.1	Early warning: Decreasing quality of call – during call *(e.)*
5.2	Early warning: Decreasing quality of call – during call *(f.)*
6.1	Fundamentally poor quality of call – during call *(e.)*
6.2	Fundamentally poor quality of call – during call *(f.)*
6.3	Fundamentally poor quality of call – before call *(e.)*
6.4	Fundamentally poor quality of call – before call *(f.)*
7. 1	Loss of connection *(e.)*
7. 2	Loss of connection *(f.)*
8.1	Active call
9. 1	Automatic handover – during call *(e.)*
9. 2	Automatic handover – during call *(f.)*
9. 3	Automatic handover – before call *(e.)*
9. 4	Automatic handover – before call *(f.)*
10.0	Menu settings: Messages

Because mobile telephone users have different preferences, the users can decide for themselves if they wish to be notified by sound, LED or vibration. Each of these elements can be set or adjusted in the application's settings menu and used in combination or individually. The user's preferences are then also reflected in the user's cost behavior: the user may prefer a less expensive connection (and a higher risk of connection loss and warnings about disruptions) or high connection quality and call continuity (with possible higher costs). The configuration menu allows for settings that reflect the preference for these service functions. Depending on the service function (costs or quality), the individual elements (sound/vibration/LED) can be set separately. In another sub-menu a choice between "optimization" and "warning" can be set according to the user's preference. Not only can the user define whether or not to be informed about cost features with an on/off switch, but a sub-section also provides an option to be warned about cost hazards or separately informed about potential cost savings.

TABLE III. EXAMPLE OF CONVENIENCE MESSAGE

9.1. Automatic handover– during call	
Element	***Description***
Graphic	Progress symbol […]
Sound	*Hold sound (starts after 2 secs. in case of delay)*
	- With successful handover: confirmation sound
	- With failed handover: error message
Text	"Please wait"
	- or with termination: e.g. "Please try to restore connection."
Vibration	-
LED	Continuous short blinking, yellow
Voice	*- Lost connection: "Please wait while we try to reconnect you."*

VI. CONCLUSION

Mobile UC/UCC telephony is a permanent service process (voice call availability and continuity). User touch-points of the UC/UCC service result from telecommunication activities. These are triggered by the user, but are controlled by permanent and situational background processes. The UC1000 mobile client including convenience messages improves the telephony experience by means of specially designed QoS sensitive messages that support the user of mobile VoIP telephony and provide greater transparency and opportunities for intervention. The increase in QoE occurs in particular through:

- Improved telephony quality resulting from the use of innovative network technology
- Automatic user navigation in handling connection problems
- Feedback about the telephone connection and call quality, and warnings about possible interruptions
- Tips for improved connection options and call quality
- Assurance of improved user telephony experience through service-oriented user interaction
- Consolidation of the QoS and QoE relationship

 ISBN: 978-1-61208-402-2

- Extension of seamless mobility

For the users of the UC1000 Mobile with HOCIS, the implementation of QoS-sensitive messages and service functions offers the potential for increased QoE in UC telephony. Beyond the basic feedback about the status of the telephony service, the UC1000 with HOCIS also offers service functions such as cost control, cost warning, quality warning, device management, and forewarning features. These features provide the user with transparency, extended options, and control of telephone bills and private and business telephony with a single device.

The merger of the extended network handover technology HOCIS with an integrated research and interaction design process through the IEB has resulted in a user-oriented expansion of the UC1000 to include the HOCIS application. Reliable connections, transparency, and user navigation are increased by means of a service-oriented interaction concept. Each user can individually configure how to be informed, warned, and assisted with QoS service functions – through text, graphics, sound, LED, vibration or voice messages. The innovative core is in the development of multisensory user navigation, which is based on an interaction concept with its own sound design, voice messages, and media elements. The result is an improved sense of QoS and increased QoE of professional UC telephony. The UC1000 with HOCIS, including QoS-sensitive convenience messages, is a system that routinely aids the professional user's telephone communication efficiently and in a service-oriented manner, across all UC functions.

This example for the implementation of convenience messaging in an existing UC client on the basis of the HOCIS technology demonstrates a unique approach for QoS-sensitive user interaction. The possibility to adapt this system to any provider's telephony infrastructure shows that the potential greatly exceeds the application in Unified Communications. Based on this result, it can be assumed that QoS-sensitive user interaction offers opportunities for wider-ranging applications in telecommunications and consolidates the QoS/QoE ratio to the user's benefit.

ACKNOWLEDGMENT

For research cooperation: Prof. Dr.-Ing. Sigram Schindler, Christian Negrutiu, Ivan Ivanovic and Marius Witke. For the temporary support of the project leadership: Dr. Anna Riedel. For input and participation in the workshops: Carl-Frank Westermann, Lenard Gimpel, Michael Pietrek, Thomas Ness and Daniel Wetzel. For discussion of the content: Jörg Müller and Jochen Maria Weber.

REFERENCES

[1] M. Emmelmann, S. Wiethoelter, A. Koepsel, C. Kappler and A. Wolisz, "Moving toward seamless mobility: state of the art and emerging aspects in standardization bodies," Wireless Personal Communications, Springer Publish On, 10/2007, DOI10.1007/s11277-007-9255-6, pp. 803-816.

[2] M. H. Shah, "Mobile working", London: Routledge, 2014, p. 7.

[3] WO 2008/064918 A2, "Handover convenience information service," [Patent], Inventors: Sigram Schindler, Doerte Schoenberg, Juergen Schulze, 05.06.2008.

[4] R. Lehmann, "Change from telecommunications to Unified Communications," Transl. [Orig.: "Wandel von der Telekommunikation zu Unified Communications",] Wiesbaden: Gabler Verlag, 2011, p. 139.

[5] R. Lehmann, "Change from telecommunications to Unified Communications," Transl. [Orig.: "Wandel von der Telekommunikation zu Unified Communications",] Wiesbaden: Gabler Verlag, 2011, p. 36.

[6] Teltarif.de Onlineverlag GmbH, "Editorial: language service, not working," Transl. [Orig.: "Editorial: Sprachdienst, kaputt,"] [Online]. Available from: http://www.teltarif.de/ausfall-sprachtelefonie-deutsche-telekom-voip/news/56856.html, [retrieved: March, 2015].

[7] Chip.de, "The hardest German mobile phone network test," Transl. [Orig.: "Der härteste Handy-Netztest Deutschlands,"] [Online]. Available from: http://www.chip.de/artikel/Der-haerteste-Handy-Netztest-Deutschlands-Telekom-Vodafone-O2-und-E-Plus-im-Test_63944005.html, [retrieved: March, 2015].

[8] IT Wissen, "CSFB (circuit switched fallback)," [Online]. Available from: http://www.itwissen.info/definition/lexikon/CSFB-circuit-switched-fallback.html, [retrieved: March, 2015].

[9] M. Richter, "Criteria of Usability," Transl. [Orig.: "Kriterien der Benutzerfreundlichkeit,"] 1997, [Online]. Available from: http://www.michaelrichter.ch/literat_97.pdf, [retrieved: March, 2015].

[10] H. Raffaseder, "Audiodesign," Munich: Hanser Verlag, 2010, p. 292.

[11] H. Raffaseder, "Audiodesign," Munich: Hanser Verlag, 2010, p. 292.

[12] H. Raffaseder, "Audiodesign," Munich: Hanser Verlag, 2010, p. 292.

[13] W. Weber (eds.), "Compendium information design," Transl. [Orig.: "Kompendium Informationsdesign,"] Berlin [u.a.]: Springer, p. 231, 2008.

[14] TELES FMC+C Innovations GmbH, "UC1000 mobile," [Online]. Available from: https://teles-fmc.com/index.php/en/products-16/uc1000-mobile, retrieved: March, 2015].

An Efficient Method for Reliability Evaluation of Data Storage Systems

Ilias Iliadis and Vinodh Venkatesan
IBM Research – Zurich
Email: {ili,ven}@zurich.ibm.com

***Abstract*—The effectiveness of the redundancy schemes that have been developed to enhance the reliability of storage systems has predominantly been evaluated based on the mean time to data loss (MTTDL) metric. This metric has been widely used to compare schemes, to assess tradeoffs, and to estimate the effect of various parameters on system reliability. Analytical expressions for MTTDL are typically derived using Markov chain models. Such derivations, however, remain a challenging task owing to the high complexity of the analysis of the Markov chains involved, and therefore the system reliability is often assessed by rough approximations. To address this issue, a general methodology based on the direct path approximation was used to obtain the MTTDL analytically for a class of redundancy schemes and for failure time distributions that also include real-world distributions, such as Weibull and gamma. The methodology, however, was developed for the case of a single direct path to data loss. This work establishes that this methodology can be extended and used in the case where there are multiple shortest paths to data loss to approximately derive the MTTDL for a broader set of redundancy schemes. The value of this simple, yet efficient methodology is demonstrated in several contexts. It is verified that the results obtained for RAID-5 and RAID-6 systems match with those obtained in previous work. As a further demonstration, we derive the exact MTTDL of a given RAID-51 system and confirm that it matches with the MTTDL obtained from the methodology proposed.**

***Keywords*–Reliability metric; MTTDL; recovery; RAID.**

I. Introduction

Storage systems experience data losses due to device failures, including disk and node failures. To avoid a permanent loss of data, redundancy schemes were developed that enable the recovery of this data. However, during rebuild operations additional device failures may occur that eventually lead to permanent data losses. There is a variety of redundancy schemes that offer different levels of reliability as they tolerate varying degrees of device failures. Each of these schemes is characterized by an overhead, which reflects the additional operations that need to be performed, and a storage efficiency, which expresses the additional amount of data, referred to as parity, that needs to be stored in the system.

The reliability of storage systems and the effectiveness of redundancy schemes have predominantly been assessed based on the mean time to data loss (MTTDL) metric, which expresses the amount of time that is expected to elapse until the first data is irrecoverably lost [1][2]. During this period, failures cause data to be temporarily lost, which is subsequently recovered owing to the redundancy built into the system.

Analytical expressions for the MTTDL are typically derived using Markov chain models [3], which assume that the times to component failures are independent and exponentially distributed. A methodology for obtaining MTTDL under general non-exponential failure and rebuild time distributions, which therefore does not involve any Markov analysis, was presented in [4]. The complexity of these derivations depends on the redundancy schemes and the underlying system configurations considered. The MTTDL metric has been proven useful for assessing tradeoffs, for comparing schemes, and for estimating the effect of various parameters on system reliability [5][6][7]. Analytical closed-form expressions for the MTTDL provide an accurate account of the effect of various parameters on system reliability. However, deriving exact closed-form expressions remains a challenging task owing to the high complexity of the analysis of the Markov chains involved [8][9]. For this reason, the system reliability is instead often assessed by rough approximations. As the direct MTTDL analysis is typically hard, an alternative is performing event-driven simulations [10][11]. However, simulations do not provide insight into how the various parameters affect the system reliability. This article addresses this issue by presenting a simple, yet efficient method, referred to as *shortest path approximation*, to obtain the MTTDL analytically for a broad set of redundancy schemes. It achieves that by considering the most likely paths that lead to data loss, which are the shortest ones. In contrast to simulations, this method provides approximate closed-form expressions for the MTTDL, thus circumventing the inherent complexity of deriving exact expressions using Markov analysis. Note also that this method was previously applied in the context of assessing system unavailability, in particular for systems characterized by large Markov chains [12]. It turns out that this approach agrees with the principle encountered in the probability context expressed by the phrase *"rare events occur in the most likely way"*. This is also demonstrated in [13], where the reliability level of systems that are highly reliable is essentially determined by the so-called "main event", which is the shortest way of failure appearance, that is, along the minimal monotone paths.

In [4][14-17], it was shown that the direct path approximation yields accurate analytical reliability results. To further investigate the validity of the direct-path-approximation method, we apply it to derive the MTTDL results for RAID-5 and RAID-6 systems and subsequently verify that they match with those obtained in previous works [1][2]. In all these previous works though, there is a single direct path to data loss. In contrast, our article is concerned with the case where there are multiple shortest paths to data loss. In this work, we investigate this issue and establish that the direct-path-

 ISBN: 978-1-61208-402-2

approximation method can be extended and also applied in the case of multiple shortest paths and yield accurate reliability results. In particular, we derive the approximate MTTDL of a RAID-51 system using the shortest path approximation. Subsequently, as a demonstration of the validity of the method proposed, we derive the exact MTTDL for a specific instance of a RAID-51 system and confirm that it matches with the corresponding MTTDL obtained using our method.

The remainder of the paper is organized as follows. Section II reviews the general framework for deriving the MTTDL of a storage system. Subsequently, the notion of the direct path to data loss is discussed in Section III, and the efficiency of the direct path approximation is demonstrated in Section IV. Section V discusses the case of multiple shortest paths to data loss and presents the analysis of a RAID-51 system. Finally, we conclude in Section VI.

II. Derivation of MTTDL

A. Markov Analysis

Continuous-time Markov chain (CTMC) models reflecting the system operation can be constructed when the device failures and rebuild times are assumed to be independent and exponentially distributed. Under these assumptions, an appropriate CTMC model can be formulated to characterize the system behavior and capture the corresponding state transitions, including those that lead to data loss. Subsequently, using the infinitesimal generator matrix approach and determining the average time spent in the transient states of the Markov chain yields a closed-form expression for the MTTDL of the system [3]. The results obtained by using CTMC models are often approximate because in practice the times to device failure and the rebuild times are not exponentially distributed. To address this issue, a more general analytical method is required.

B. Non-Markov Analysis

Here we briefly review the general framework for deriving the MTTDL developed in [4][14] using an analytical approach that does not involve any Markov analysis, and therefore avoids the deficiencies of Markov models. The underlying models are not semi-Markov, in that the the system evolution does not depend only on the latest state, but also on the entire path that led to that state. In particular, it depends on the fractions of the data not rebuilt when entering each state. In [18] it was demonstrated that a careless evaluation of these fractions may in fact easily lead to erroneous results.

At any point of time, the system can be thought to be in one of two modes: normal mode and rebuild mode. During normal mode, all data in the system has the original amount of redundancy and there is no active rebuild in process. During rebuild mode, some data in the system has less than the original amount of redundancy and there is an active rebuild process that is trying to restore the redundancy lost. A transition from normal to rebuild mode occurs when a device fails; we refer to the device failure that causes this transition as a *first-device* failure. Following a first-device failure, a complex sequence of rebuild operations and subsequent device failures may occur, which eventually leads the system either to an irrecoverable data loss (DL), with the probability of this event denoted by P_{DL}, or back to the original normal mode by restoring all replicas lost. Typically, the rebuild times are much shorter than the times to failure. Consequently, the time required for this complex sequence of events to complete is negligible compared with the time between successive first-device failures, and therefore can be ignored.

Let T_i be the ith interval of a fully operational period, that is, the time interval from the time t that the system is brought to its original state until a subsequent first-device failure occurs. As the system becomes stationary, the length of T_i converges to T. In particular, for a system comprising N devices with a mean time to failure of a device equal to $1/\lambda$, the expected length of T is given by [4]

$$E(T) := \lim_{i\to\infty} E(T_i) = 1/(N\,\lambda)\ . \tag{1}$$

The notation used is given in Table I. Note that the methodology presented here does not involve any Markov analysis and holds for general failure time distributions, which can be exponential or non-exponential, such as the Weibull and gamma distributions.

As the probability that each first-device failure results in data loss is P_{DL}, the expected number of first-device failures until data loss occurs is $1/P_{\mathrm{DL}}$. Thus, by neglecting the effect of the relatively short transient rebuild periods of the system, the MTTDL is essentially the product of the expected time between two first-device-failure events, $E(T)$, and the expected number of first-device-failure events, $1/P_{\mathrm{DL}}$:

$$\text{MTTDL} \approx \frac{E(T)}{P_{\mathrm{DL}}}\ . \tag{2}$$

Substituting (1) into (2) yields

$$\text{MTTDL} \approx \frac{1}{N\,\lambda\,P_{\mathrm{DL}}}\ . \tag{3}$$

III. Direct Path to Data Loss

As mentioned in Section II, during rebuild mode, some data in the system has less than the original amount of redundancy and there is an active rebuild process that aims at restoring the lost redundancy. The direct path to data loss represents the most likely scenario that leads to data loss. This path considers the smallest number of subsequent device failures that occur while the system is in rebuild mode and lead to data loss.

The direct-path-approximation method was applied in [4][14] and led to an analytical approach that does not involve any Markov analysis, and therefore avoids the deficiencies of Markov models. This approach yields accurate results when the storage devices are highly reliable, that is, when the ratio of the mean rebuild time $1/\mu$ (typically on the order of tens of hours) to the mean time to failure of a device $1/\lambda$ (typically on the order of a few years) is very small:

$$\frac{1}{\mu} \ll \frac{1}{\lambda}\ , \quad \text{or} \quad \frac{\lambda}{\mu} \ll 1\ , \quad \text{or} \quad \lambda \ll \mu\ . \tag{4}$$

TABLE I. Notation of System Parameters

Parameter	Definition
N	Number of devices in the system or in an array group
$1/\lambda$	Mean time to failure for a device
$1/\mu$	Mean time to rebuild

More specifically, this approach considers the system to be in exposure level e when the maximum number of replicas lost by each of the data is equal to e. Let us consider, for instance, a replication-based storage system, where user data is replicated r times. In this case, the system is in exposure level e if there exists data with $r - e$ copies, but there is no data with fewer than $r - e$ copies. Device failures and rebuild processes cause the exposure level to vary over time. Consider the direct path of successive transitions from exposure level 1 to r. In [14], it was shown that P_{DL} can be approximated by the probability of the direct path to data loss, $P_{\text{DL,direct}}$, when devices are highly reliable, that is,

$$P_{\text{DL}} \approx P_{\text{DL,direct}} = \prod_{e=1}^{r-1} P_{e \to e+1} , \tag{5}$$

where $P_{e \to e+1}$ denotes the transition probability from exposure level e to $e + 1$. In fact, the above approximation holds for arbitrary device failure time distributions, and the relative error tends to zero as for highly reliable devices the ratio λ/μ tends to zero [4]. The MTTDL is then obtained by substituting (5) into (3). In [16], the direct path methodology is extended to more general erasure codes, which include RAID systems.

Note that this analysis can also be applied to assess reliability, in terms of the MTTDL, for systems modeled using a CTMC. For instance, in [5], a RAID-5 system that was modeled using a CTMC was analyzed by both a Markov analysis and an approach similar to the general framework. This fact is used in Section IV to compare the MTTDL of RAID systems obtained using the direct path approximation in the context of the general framework with the corresponding MTTDL obtained using Markov analysis of CTMCs. This approach is simpler, in that it circumvents the inherent complexity of deriving exact MTTDL expressions using Markov analysis. In Section V, we demonstrate that the direct-path-approximation method can also be extended and applied in the case of multiple shortest paths. We establish this for a system modeled using a CTMC, and conjecture that this should also hold in the case of non-Markovian systems.

IV. Comparison of Markov Analysis and Direct Path Approximation

A common scheme used for tolerating device (disk) failures is the redundant array of independent disks (RAID) [1][2]. The RAID-5 scheme arranges devices in groups (arrays), each with one redundant device, and can tolerate one device failure per array. Similarly, the RAID-6 scheme arranges devices in arrays, each with two redundant devices, and can tolerate up to two device failures per array. Considering that an array consists of N devices, the storage efficiency of a RAID-5 and RAID-6 system is $(N - 1)/N$ and $(N - 2)/N$, respectively.

It turns out that the MTTDL of systems comprised of highly reliable devices can be approximated by using the *direct path approximation*. We proceed to demonstrate this by presenting two specific examples, the RAID-5 and RAID-6 systems. In both cases, the RAID array is assumed to contain N devices, and the numbered states of the corresponding Markov models represent the number of failed devices. The DL state represents a data loss due to a device failure that occurs when the system is in the critical mode of operation. A RAID array is considered to be in *critical mode* when an additional device failure can no longer be tolerated. Thus, RAID-5 and RAID-6 arrays are in critical mode when there are $N - 1$ devices and $N - 2$ devices in operation, that is, when they operate with one device and two devices failed, respectively.

A. MTTDL *for a RAID-5 Array*

The Markov chain model for a RAID-5 array is shown in Fig. 1. When the first device fails, the array enters critical mode, which corresponds to the transition from state 0 to state 1. As initially there are N devices in operation, the mean time until the first failure is equal to $1/(N\lambda)$, and the corresponding transition rate is its inverse, that is, $N\lambda$. Subsequently, the critical mode ends owing to either a successful completion of the rebuild or another device failure. The former event is represented by the state transition from state 1 to state 0 with a rate of μ, given that the mean rebuild time is equal to $1/\mu$. The latter event leads to data loss and is represented by the state transition from state 1 to state DL with a rate of $(N-1)\lambda$ given that in critical mode there are $N-1$ devices in operation.

The exact MTTDL, denoted by $\text{MTTDL}_{\text{RAID-5}}$, is obtained from [5, Eq. (45)] by setting $P_{\text{uf}}^{(1)} = 0$ as follows:

$$\text{MTTDL}_{\text{RAID-5}} = \frac{\mu + (2N-1)\lambda}{N\,(N-1)\,\lambda^2} \,. \tag{6}$$

Note that when $\lambda \ll \mu$, the first term of the numerator in (6) can be ignored, such that the $\text{MTTDL}_{\text{RAID-5}}$ can be approximated by $\text{MTTDL}_{\text{RAID-5}}^{(\text{approx})}$ as follows:

$$\text{MTTDL}_{\text{RAID-5}}^{(\text{approx})} \approx \frac{\mu}{N(N-1)\,\lambda^2} \,. \tag{7}$$

This result was obtained in [1] by using an approach that is essentially the direct path approximation. Next, we present this derivation for completeness. The transition from state 0 to state 1 represents the first device failure. The direct path to data loss involves a subsequent device failure prior to completing the rebuild process and returning to state 0. This corresponds to the state transition from state 1 to state DL, with the corresponding probability $P_{1 \to DL}$ given by

$$P_{\text{DL}} = P_{\text{DL,direct}} = P_{1 \to DL} = \frac{(N-1)\,\lambda}{\mu + (N-1)\,\lambda} \,. \tag{8}$$

Substituting (8) into (3) yields

$$\text{MTTDL}'_{\text{RAID-5}} \approx \frac{\mu + (N-1)\,\lambda}{N(N-1)\,\lambda^2} \,. \tag{9}$$

Note that the approximation given in (7) now follows immediately from (9) by using (4) and therefore neglecting the second term of the numerator.

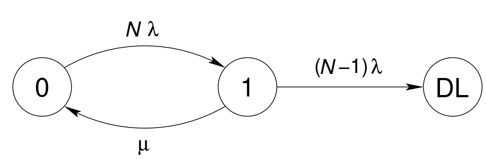

Figure 1. Reliability model for a RAID-5 array.

B. MTTDL *for a RAID-6 Array*

The Markov chain model for a RAID-6 array is shown in Fig. 2. The first device failure is represented by the transition from state 0 to state 1. As initially there are N devices in operation, the mean time until the first failure is $1/(N\lambda)$, and the corresponding transition rate is its inverse, that is, $N\lambda$. The system exits from state 1 owing to either a successful completion of the rebuild or another device failure. The former event is represented by the state transition from state 1 to state 0 with a rate of μ. The latter event is represented by the state transition from state 1 to state 2 with a rate of $(N-1)\lambda$. Subsequently, the system exits from state 2 owing to either a successful completion of the rebuild or another device failure. The former event is represented by the state transition from state 2 to state 0 with a rate of μ, given that the mean rebuild time is equal to $1/\mu$. The latter event leads to data loss and is represented by the state transition from state 2 to state DL with a rate of $(N-2)\lambda$ given that in critical mode there are $N-2$ devices in operation.

The exact MTTDL, denoted by $\text{MTTDL}_{\text{RAID-6}}$, is obtained from [5, Eq. (45)] by setting $\mu_1 = \mu_2 = \mu$ and $P_{\text{uf}}^{(\text{r})} = P_{\text{uf}}^{(2)} = 0$ as follows:

$$\text{MTTDL}_{\text{RAID-6}} = \frac{\mu^2 + 3(N-1)\lambda\mu + (3N^2 - 6N + 2)\,\lambda^2}{N\,(N-1)\,(N-2)\,\lambda^3} . \tag{10}$$

Note that when $\lambda \ll \mu$, the last two terms of the numerator of (10) can be neglected and thus $\text{MTTDL}_{\text{RAID-6}}$ can be approximated by $\text{MTTDL}_{\text{RAID-6}}^{(\text{approx})}$ as follows:

$$\text{MTTDL}_{\text{RAID-6}}^{(\text{approx})} \approx \frac{\mu^2}{N(N-1)(N-2)\,\lambda^3} , \tag{11}$$

which is the same result as that reported in [2].

We now proceed to show how the approximate MTTDL of the system can be derived in a straightforward manner by applying the direct-path-approximation technique. The transition from state 0 to state 1 represents the first device failure. The direct path to data loss involves two subsequent device failures prior to completing the rebuild process and returning to state 0. This corresponds to the state transitions from state 1 to state 2 and from state 2 to state DL, with the corresponding probabilities $P_{1\to 2}$ and $P_{2\to DL}$ given by

$$P_{1\to 2} = \frac{(N-1)\,\lambda}{\mu + (N-1)\,\lambda} . \tag{12}$$

and

$$P_{2\to DL} = \frac{(N-2)\,\lambda}{\mu + (N-2)\,\lambda} . \tag{13}$$

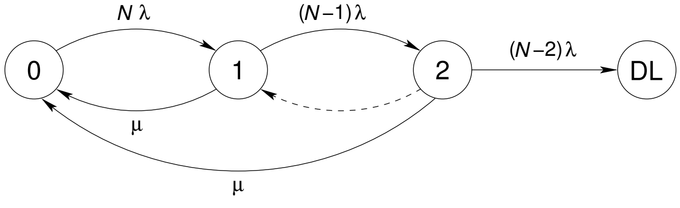

Figure 2. Reliability model for a RAID-6 array.

Thus, the probability of data loss, that is, the probability that from state 1 the system goes to state DL prior to reaching state 0, is equal to

$$\begin{aligned} P_{\text{DL}} = P_{\text{DL,direct}} &= P_{1\to 2}\, P_{2\to DL} \\ &= \frac{(N-1)\,\lambda}{\mu + (N-1)\,\lambda} \cdot \frac{(N-2)\,\lambda}{\mu + (N-2)\,\lambda} \qquad (14) \\ &\approx \frac{(N-1)(N-2)\,\lambda^2}{\mu^2} , \qquad (15) \end{aligned}$$

where the approximation is obtained by using (4) and therefore neglecting the second terms of the denominators in (14).

We now verify that substituting (15) into (3) yields the approximation given in (11).

Remark 1. If the transition from state 2 to state 0 was not to state 0, but was instead to state 1, as shown in Fig. 2 by the dashed arrow, the expression for $P_{2\to DL}$ given by (13) would still hold. However, in this case it would hold that $P_{\text{DL}} > P_{\text{DL,direct}}$ as, in addition to the direct path $1 \to 2 \to \text{DL}$, there are other possible paths $1 \to 2 \to 1 \to 2 \to \cdots \to 1 \to 2 \to \text{DL}$ to data loss. In [14] it was shown that, for highly reliable systems, the direct path dominates the effect of all other possible paths and therefore its probability, $P_{\text{DL,direct}}$, approximates well the probability of all paths, P_{DL}, that is,

$$P_{\text{DL}} \approx P_{\text{DL,direct}} = P_{1\to 2}\, P_{2\to DL} \approx \frac{(N-1)(N-2)\,\lambda^2}{\mu^2} . \tag{16}$$

In this case, the MTTDL is given by

$$\text{MTTDL}'_{\text{RAID-6}} = \frac{(3N^2 - 6N + 2)\,\lambda^2 + 2(N-1)\lambda\mu + \mu^2}{N\,(N-1)\,(N-2)\,\lambda^3} , \tag{17}$$

which, as expected, is less than that given in (10). Despite this difference, the approximation given in (11) still holds because (16) is the same as (15),.

V. Multiple Shortest Paths to Data Loss

We now consider redundancy schemes for which there are multiple shortest paths to data loss. Following the analysis presented in [14] for the direct path approximation, we conjecture that, for highly reliable systems, the shortest paths dominate the effect of all other possible paths and therefore the sum of their corresponding probabilities, $P_{\text{DL,shortest}}$, approximates well the probability of all paths, P_{DL}, that is,

$$P_{\text{DL}} \approx P_{\text{DL,shortest}} . \tag{18}$$

A. *A RAID-51 Array*

We proceed by considering a RAID-51 system, which is a RAID-5 array with mirroring. The contents of failed devices are recovered by their mirrors, and if this is not possible, they are recovered through the corresponding RAID-5 arrays. The configuration comprises D pairs of mirrored devices, where each pair contains two devices with identical content. It therefore consists of two identical RAID-5 arrays, for a total of $N\,(= 2D)$ devices. This configuration was considered in [9], referred to as RAID 5+1, with the corresponding Markov model shown in [9, Fig. 7(a)]. It is redrawn in Fig. 3 with the parameters λ and μ corresponding to the parameters μ and ν of the initial figure, respectively. Also, the DL states correspond

 ISBN: 978-1-61208-402-2

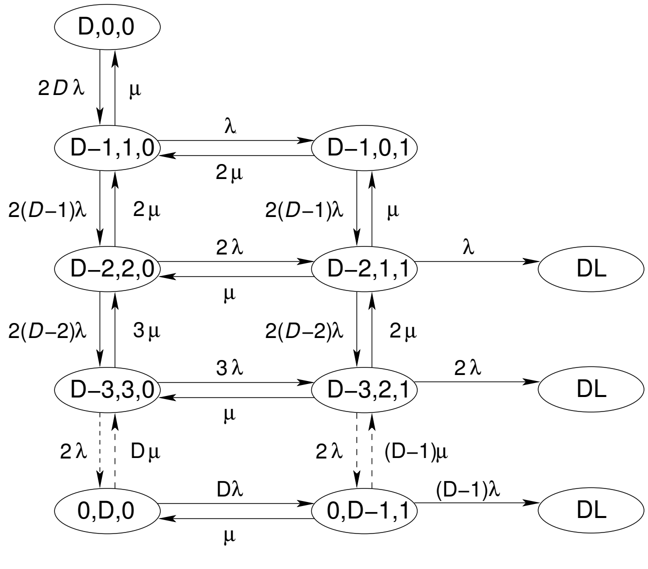

Figure 3. Reliability model for a RAID-51 array.

to the 'Failure' states, and the state tuples (x, y, z) indicate that there are x pairs with both devices in operation, y pairs with one device in operation and one device failed, and z pairs with both devices failed. Also, some typos regarding the transition rates were corrected.

An exact evaluation of the MTTDL associated with this Markov chain model appears to be a very challenging, if not infeasible, task. Thus, in [9] a rough approximation was obtained by first deriving the failure and repair rates for a mirrored pair of devices, and then substituting these values into expression (7) for a single RAID-5 system. The MTTDL is obtained in [9, Eq. (11)] as follows:

$$\text{MTTDL} \approx \frac{\mu^3}{4D(D-1)\,\lambda^4} \,. \tag{19}$$

B. MTTDL *Evaluation Using the Shortest Path Approximation*

The transition from $(D, 0, 0)$ to state $(D-1, 1, 0)$ represents the first device failure. As initially there are $2\,D$ devices in operation, the mean time until the first failure is $1/(2\,D\lambda)$, and the corresponding transition rate is its inverse, $2\,D\lambda$.

The most likely path to data loss is the shortest path from state $(D-1, 1, 0)$ to a DL state, which in this case comprises two such paths, as shown in Fig. 4: the upper path $(D-1, 1, 0) \to (D-1, 0, 1) \to (D-2, 1, 1) \to$ DL and the lower path: $(D-1, 1, 0) \to (D-2, 2, 0) \to (D-2, 1, 1) \to$ DL. Each of these paths involves three subsequent device failures.

After the first device has failed, there are $D-1$ pairs with both devices in operation, and one pair, say PR_1, with one device in operation and one device failed, which corresponds to the transition from state $(D, 0, 0)$ to state $(D-1, 1, 0)$. The rebuild of the failed device consists of recovering its data to a new spare device by copying the contents of its mirror to it, that is, of the device in operation in PR_1. Then, the next event can be either a successful completion of the rebuild or another device failure. The former event is represented by the

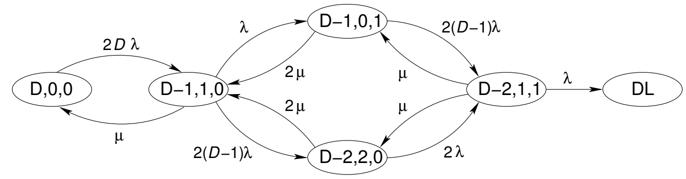

Figure 4. Shortest-path reliability model for a RAID-51 array.

state transition from state $(D-1, 1, 0)$ to state $(D, 0, 0)$ with a rate of μ. For the latter event, two cases are considered:

Case a) The second device that fails is the device in operation concerning pair PR_1, which corresponds to the transition from state $(D-1, 1, 0)$ to state $(D-1, 0, 1)$, as both devices of pair PR_1 have failed, and all other $D-1$ pairs remain intact. The transition rate is λ, which is the failure rate of the last failed device. The contents of the devices of pair PR_1 are recovered through the corresponding RAID-5 arrays. As both devices of pair PR_1 are under rebuild, the transition rate from state $(D-1, 0, 1)$ back to state $(D-1, 1, 0)$ is $2\,\mu$. If, however, prior to the completion of any of the two rebuilds another device of the remaining $2\,(D-1)$ fails, then there will be $D-2$ pairs with both devices in operation, one pair, say PR_2, with one device in operation and one device failed, and pair PR_1 with both devices failed. This corresponds to the transition from state $(D-1, 0, 1)$ to state $(D-2, 1, 1)$ with a corresponding transition rate equal to $2\,(D-1)\,\lambda$. Note that in [9, Fig. 7(a)] this transition rate is erroneously indicated as $(2\,D-1)\,\mu$ instead of $2\,(D-1)\,\mu$.

Case b) The second device that fails is one of the $2\,(D-1)$ devices in the $D-1$ pairs, say a device concerning PR_2. This corresponds to the transition from state $(D-1, 1, 0)$ to state $(D-2, 2, 0)$, as both pairs PR_1 and PR_2 have one device in operation and one device failed, and all other $D-2$ pairs remain intact. The corresponding transition rate is equal to $2\,(D-1)\,\lambda$. Note that in [9, Fig. 7(a)] this transition rate is erroneously indicated as $(2\,D-1)\,\mu$ instead of $2\,(D-1)\,\mu$. The contents of the failed devices are recovered from their corresponding mirrors. As both devices of the two pairs PR_1 and PR_2 are under rebuild, the transition rate from state $(D-2, 2, 0)$ back to state $(D-1, 1, 0)$ is $2\,\mu$. If, however, prior to the completion of any of the two rebuilds another device of the two remaining devices in operation in PR_1 and PR_2 fails (say, that of pair PR_1), then there will be $D-2$ pairs with both devices in operation, one pair (PR_2) with one device in operation and one device failed, and one pair (PR_1) with both devices failed. This corresponds to the transition from state $(D-2, 2, 0)$ to state $(D-2, 1, 1)$, with a corresponding transition rate $2\,\lambda$.

At state $(D-2, 1, 1)$, the device in pair PR_2 failed is recovered by its mirror. However, the corresponding failed device in pair PR_1 cannot be recovered because the RAID-5 array has suffered two device failures. In contrast, the other failed device in pair PR_1 can by recovered because the corresponding RAID-5 array has suffered only one device failure.

The completion of the rebuild of the failed device in pair PR_2 corresponds to the transition from state $(D-2, 1, 1)$ to state $(D-1, 0, 1)$, with a transition rate of μ. The completion of the rebuild of the failed device in pair PR_1 through the RAID capability corresponds to the transition from state $(D-2,1,1)$ to state $(D-2,2,0)$, with a transition rate of μ. Note that in [9, Fig. 7(a)] this transition rate is erroneously indicated as $2\,\mu$ instead of μ. If, however, prior to the completion of any of these rebuilds, the device still in operation of pair PR_2 fails, this leads to data loss, as there will be two pairs failed, with each of the RAID-5 arrays having two devices failed. This corresponds to the transition from state $(D-2,1,1)$ to state DL, with a corresponding rate of λ.

The probabilities of the transitions discussed above are given by

$$P_{(D-1,1,0)\to(D-1,0,1)} = \frac{\lambda}{\mu + (2D-1)\,\lambda}\,, \tag{20}$$

$$P_{(D-1,0,1)\to(D-2,1,1)} = \frac{2(D-1)\,\lambda}{2\,\mu + 2(D-1)\,\lambda}\,, \tag{21}$$

$$P_{(D-1,1,0)\to(D-2,2,0)} = \frac{2(D-1)\,\lambda}{\mu + (2D-1)\,\lambda}\,, \tag{22}$$

$$P_{(D-2,2,0)\to(D-2,1,1)} = \frac{2\,\lambda}{2\,\mu + 2\,\lambda}\,, \tag{23}$$

and

$$P_{(D-2,1,1)\to\mathrm{DL}} = \frac{\lambda}{2\,\mu + \lambda}\,. \tag{24}$$

Consequently, the probability of the upper path to data loss, P_u, is given by

$$\begin{aligned} P_u =& P_{(D-1,1,0)\to(D-1,0,1)}\, P_{(D-1,0,1)\to(D-2,1,1)}\, P_{(D-2,1,1)\to\mathrm{DL}} \\ =& \frac{\lambda}{\mu + (2D-1)\,\lambda} \cdot \frac{2(D-1)\,\lambda}{2\,\mu + 2(D-1)\,\lambda} \cdot \frac{\lambda}{2\,\mu+\lambda}\,, \end{aligned} \tag{25}$$

and that of the lower path to data loss, P_l, is given by

$$\begin{aligned} P_l =& P_{(D-1,1,0)\to(D-2,2,0)}\, P_{(D-2,2,0)\to(D-2,1,1)}\, P_{(D-2,1,1)\to\mathrm{DL}} \\ =& \frac{2(D-1)\,\lambda}{\mu + (2D-1)\,\lambda} \cdot \frac{2\,\lambda}{2\,\mu + 2\,\lambda} \cdot \frac{\lambda}{2\,\mu+\lambda}\,, \end{aligned} \tag{26}$$

By considering (4), (25) and (26) yield the following approximations:

$$P_u \;\approx\; \frac{\lambda}{\mu} \cdot \frac{2(D-1)\,\lambda}{2\,\mu} \cdot \frac{\lambda}{2\,\mu} \;=\; \frac{(D-1)\,\lambda^3}{2\mu^3} \tag{27}$$

and

$$P_l \;\approx\; \frac{2(D-1)\,\lambda}{\mu} \cdot \frac{\lambda}{\mu} \cdot \frac{\lambda}{2\,\mu} \;=\; \frac{(D-1)\,\lambda^3}{\mu^3}\,. \tag{28}$$

The probability of the shortest paths to data loss, $P_{\mathrm{DL,shortest}}$, is the sum of P_u and P_l, which by using (18), (27), and (28), yields

$$P_{\mathrm{DL}} \;\approx\; P_{\mathrm{DL,shortest}} = P_u + P_l \;\approx\; \frac{3(D-1)\,\lambda^3}{2\mu^3}\,. \tag{29}$$

Substituting (29) into (3), and considering $N = 2D$, yields the approximate MTTDL of the RAID-51 system, $\mathrm{MTTDL}^{(\mathrm{approx})}_{\text{RAID-51}}$ given by

$$\mathrm{MTTDL}^{(\mathrm{approx})}_{\text{RAID-51}} \;\approx\; \frac{\mu^3}{3D(D-1)\,\lambda^4}\,. \tag{30}$$

Remark 2. Note that the prediction given by (30) is higher than that obtained in [9], which is given by (19). At first glance, this seems to be counterintuitive. The approximation in [9] considers only failures of mirrored pair of devices, which corresponds to the upper path to data loss. As this neglects the lower path, one would expect the prediction in [9] to be higher, not lower. The reason for this counterintuitive result is the fact that considering additional paths on the one hand may increase the number of paths that lead to data loss, but on the other hand it may also increase the number of the paths that do not lead to data loss, therefore delaying the occurrence of data loss. For instance, when the lower path is neglected, the probability $P_{(D-2,1,1)\to\mathrm{DL}}$ of the transition from state $(D-2,1,1)$ to state DL is equal to $\lambda/(\lambda+\mu)$, which is greater than the corresponding one given by (24), if also the lower path is considered.

C. Exact MTTDL *Evaluation for D=3*

An exact evaluation of the reliability of a RAID-51 system through the MTTDL associated with the corresponding Markov chain model shown in Fig. 3 appears to be a very challenging, if not infeasible, task for arbitrary D. We therefore proceed by considering a RAID-51 system with $D = 3$. The corresponding Markov chain model is shown in Fig. 5. The exact MTTDL of this system, denoted by $\mathrm{MTTDL}^{(D=3)}_{\text{RAID-51}}$, is obtained by using the infinitesimal generator matrix approach and determining the average time spent in the transient states of the Markov chain [3] Because of space limitations, we only provide the final result:

$$\mathrm{MTTDL}^{(D=3)}_{\text{RAID-51}} = \frac{2 + 20\frac{\lambda}{\mu} + 93(\frac{\lambda}{\mu})^2 + 287(\frac{\lambda}{\mu})^3 + 677(\frac{\lambda}{\mu})^4 + 939(\frac{\lambda}{\mu})^5 + 630(\frac{\lambda}{\mu})^6}{12\,\lambda^4\,\mu^{-3}\,[3 + 18\frac{\lambda}{\mu} + 35(\frac{\lambda}{\mu})^2 + 30(\frac{\lambda}{\mu})^3]}\,. \tag{31}$$

Note that when $\lambda \ll \mu$, $\mathrm{MTTDL}^{(D=3)}_{\text{RAID-51}}$ can be approximated by $\mathrm{MTTDL}^{(D=3,\mathrm{approx})}_{\text{RAID-51}}$ as follows:

$$\mathrm{MTTDL}^{(D=3,\mathrm{approx})}_{\text{RAID-51}} \;\approx\; \frac{\mu^3}{18\,\lambda^4}\,, \tag{32}$$

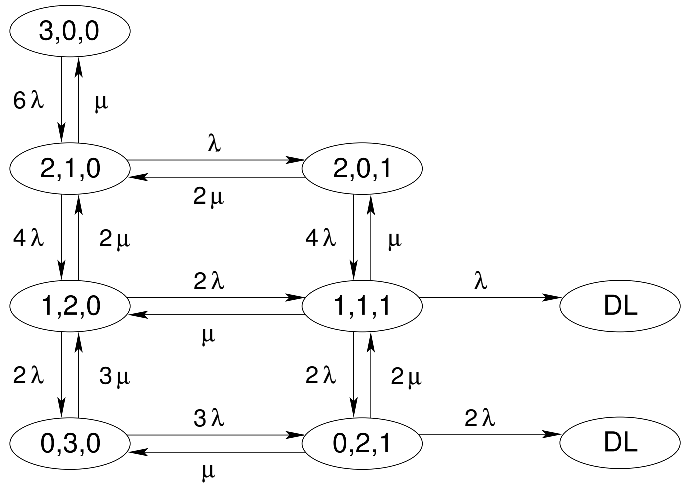

Figure 5. Reliability model for a RAID-51 array with $D = 3$.

 ISBN: 978-1-61208-402-2

which is the same result as that predicted by (30) for $D = 3$, and therefore confirms its validity.

VI. Conclusions

We considered the mean time to data loss (MTTDL) metric, which assesses the reliability level of storage systems. This work presented a simple, yet efficient methodology to approximately assess it analytically for highly reliable systems and a broad set of redundancy schemes. We extended the direct path approximation to a more general method that considers all shortest paths that lead to data loss. We subsequently applied this method to obtain a closed-form expression for the MTTDL of a RAID-51 system. We also considered a specific instance of a RAID-51 system, then derived the corresponding exact MTTDL, and subsequently confirmed that it matches that obtained from the shortest-path-approximation method. As the direct path approximation accurately predicts the reliability of non-Markovian systems with a single shortest path, we conjecture that the shortest-path-approximation method would also accurately predict the reliability of non-Markovian systems with multiple shortest paths.

Application of the shortest-path-approximation methodology developed to derive the MTTDL for systems using other redundancy schemes, such as erasure codes, is a subject of future work.

References

[1] D. A. Patterson, G. Gibson, and R. H. Katz, “A case for redundant arrays of inexpensive disks (RAID),” in Proceedings of the ACM SIGMOD International Conference on Management of Data, Jun. 1988, pp. 109–116.

[2] P. M. Chen, E. K. Lee, G. A. Gibson, R. H. Katz, and D. A. Patterson, “RAID: High-performance, reliable secondary storage,” ACM Comput. Surv., vol. 26, no. 2, Jun. 1994, pp. 145–185.

[3] K. S. Trivedi, Probabilistic and Statistics with Reliability, Queueing and Computer Science Applications, 2nd ed. New York: Wiley, 2002.

[4] V. Venkatesan and I. Iliadis, “A general reliability model for data storage systems,” in Proceedings of the 9th International Conference on Quantitative Evaluation of Systems (QEST), Sep. 2012, pp. 209–219.

[5] A. Dholakia, E. Eleftheriou, X.-Y. Hu, I. Iliadis, J. Menon, and K. Rao, “A new intra-disk redundancy scheme for high-reliability RAID storage systems in the presence of unrecoverable errors,” ACM Trans. Storage, vol. 4, no. 1, May 2008, pp. 1–42.

[6] A. Thomasian and M. Blaum, “Higher reliability redundant disk arrays: Organization, operation, and coding,” ACM Trans. Storage, vol. 5, no. 3, Nov. 2009, pp. 1–59.

[7] I. Iliadis, R. Haas, X.-Y. Hu, and E. Eleftheriou, “Disk scrubbing versus intradisk redundancy for RAID storage systems,” ACM Trans. Storage, vol. 7, no. 2, Jul. 2011, pp. 1–42.

[8] K. Rao, J. L. Hafner, and R. A. Golding, “Reliability for networked storage nodes,” IEEE Trans. Dependable Secure Comput., vol. 8, no. 3, May 2011, pp. 404–418.

[9] Q. Xin, E. L. Miller, T. J. E. Schwarz, D. D. E. Long, S. A. Brandt, and W. Litwin, “Reliability mechanisms for very large storage systems,” in Proceedings of the 20th IEEE/11th NASA Goddard Conference on Mass Storage Systems and Technologies (MSST), Apr. 2003, pp. 146–156.

[10] Q. Xin, T. J. E. Schwarz, and E. L. Miller, “Disk infant mortality in large storage systems,” in Proceedings of the 13th Annual IEEE/ACM International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS), Sep. 2005, pp. 125–134.

[11] A. Wildani, T. J. E. Schwarz, E. L. Miller, and D. D. E. Long, “Protecting against rare event failures in archival systems,” in Proceedings of the 17th Annual IEEE/ACM International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS), Sep. 2009, pp. 1–11.

[12] M. Bouissou and Y. Lefebvre, “A path-based algorithm to evaluate asymptotic unavailability for large markov models,” in Proceedings of the 48th Annual Reliability and Maintainability Symposium, Jan. 2002, pp. 32–39.

[13] I. B. Gertsbakh, “Asymptotic methods in reliability theory: A review,” Advances in Applied Probability, vol. 16, no. 1, Mar. 1984, pp. 147–175.

[14] V. Venkatesan, I. Iliadis, C. Fragouli, and R. Urbanke, “Reliability of clustered vs. declustered replica placement in data storage systems,” in Proceedings of the 19th Annual IEEE/ACM International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS), Jul. 2011, pp. 307–317.

[15] V. Venkatesan, I. Iliadis, and R. Haas, “Reliability of data storage systems under network rebuild bandwidth constraints,” in Proceedings of the 20th Annual IEEE International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS), Aug. 2012, pp. 189–197.

[16] V. Venkatesan and I. Iliadis, “Effect of codeword placement on the reliability of erasure coded data storage systems,” in Proceedings of the 10th International Conference on Quantitative Evaluation of Systems (QEST), Sep. 2013, pp. 241–257.

[17] ——, “Effect of latent errors on the reliability of data storage systems,” in Proceedings of the 21th Annual IEEE International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS), Aug. 2013, pp. 293–297.

[18] I. Iliadis and V. Venkatesan, “Expected annual fraction of data loss as a metric for data storage reliability,” in Proceedings of the 22nd Annual IEEE International Symposium on Modeling, Analysis, and Simulation of Computer and Telecommunication Systems (MASCOTS), Sep. 2014, pp. 375–384.

 ISBN: 978-1-61208-402-2

The min-max in LC Versus the max-log MAP in LC Method for Soft-Decision Decoding of MTR Codes

Nikola Djuric, Vojin Senk
Faculty of Technical Sciences, University of Novi Sad
Novi Sad, Serbia
E-mail: ndjuric@uns.ac.rs, vojin_senk@uns.ac.rs

***Abstract* – The approaches for soft-decision decoding of maximum-transition-run (MTR) codes have been extensively researched lately. Two different methods, named as *min-max in LC* and *max-log MAP in LC* have emerged as a fairly suitable for MTR code utilization in magnetic recording systems. Both methods were designed to enable the propagation of the soft-values through the logic circuits of the MTR decoder, allowing the concatenation of MTR code with some powerful error-correcting codes, such as the low-density parity-check (LDPC) codes. In this paper, the *min-max in LC* versus the *max-log MAP in LC* approach is considered in the framework of simple and the straightforward LDPC-MTR concatenation over E^2PR4 magnetic recording channel. Both methods show nearly the same coding gain of 2 dB at BER = 10^{-5}.**

Keywords – soft-decision decoding; constrained coding; MTR codes; LDPC codes.

I. Introduction

The maximum-transition-run (MTR) codes have been confirmed as very useful constrained codes for E^2PR4 magnetic recording channel, preventing the dominant error-event ± [+1 –1 +1] and consequently improving the distance properties of the channel by increasing the minimal squared Euclidean distance between two distinct sequences in this type of the channel. Their design was primarily oriented towards constrained features, resulting with unsatisfactory error correcting capabilities. Thus, the MTR codes need to be merged with some powerful error-correcting code, such as the well-established low-density parity-check (LDPC) code [2].

There were several attempts to combine/concatenate the MTR and LDPC [3]-[5], requiring that the MTR decoder is able to handle soft-values [6]. Independently, an approach of a simple and straightforward serial concatenation of these codes was considered, providing as a result two additional methods for the soft-decision decoding of MTR codes: the *min-max in LC* and *max-log MAP in LC* methods [7], [8].

In this paper, the *min-max in LC* and *max-log MAP in LC* methods are considered and compared as methods that enable propagation of the soft-values through the logic circuits of the MTR decoder in LDPC-MTR concatenation, where LDPC acts as an outer and MTR as an inner code, over the one-track one-head E^2PR4 magnetic recording channel.

Section II presents a brief overview of the *min-max in LC* and *max-log MAP in LC* approaches, while Section III explains the LDPC-MTR encoding over the E^2PR4 channel. Section IV brings complexity analyses of the two proposed methods and Section V offers simulation results and performance comparison. Finally, Section VI presents the conclusion to this paper.

II. MTR Codes and Soft-decision Decoding

The MTR codes bring a constraint in the encoding process by eliminating three or more consecutive transitions in the recording sequence of E^2PR4 channel, and thus eliminate the most troubleshooting error-event ± [+1 –1 +1], enhancing E^2PR4 channel distance properties [10].

The MTR codes were originally presented by Moon and Brickner [1]. One of their major benefits is that they can be realized using simple logic circuits, as it is presented in Figure 1, for the well know rate 4/5 (2, k = 8) MTR code.

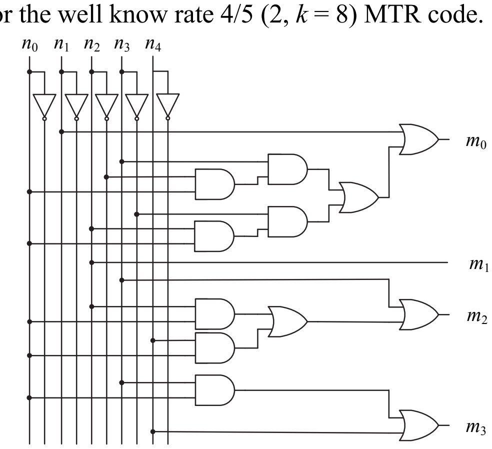

Figure 1. Rate 4/5 (2, k = 8) MTR decoder.

Such feature enables an easy hardware implementation, providing the low-cost realization for the MTR decoder.

The key idea of the *min-max in LC* and the *max-log MAP in LC* methods was to implement the soft-decision approach in existing logic circuits and to assemble the novel MTR decoder with soft-decision capabilities. The soft-value of the variable x is defined as

$$L(x) = \log[P(x=1) / P(x=0)]\,, \qquad (1)$$

where *log*() function is a natural logarithm. The sign of $L(x)$ represents the binary decision of a variable, the so called hard-decision, while the magnitude represents the confidence or reliability of a binary decision [6].

Using the soft-decision approach, the MTR decoder will be able to offer the binary decision, as well as the reliability of that decision for all its outputs, increasing information redundancy that will be delivered to the next step decoder. Such new MTR decoder that is capable to handle input soft-values and produce subsequent soft-values on its outputs is necessary in iterative decoding schemes.

A. Brief overview of the min-max in LC method

The *min-max in LC* method suggests enhancing the existing Boolean logic circuits, so that they are able to produce the output soft-values according to the following expressions [7], [11]:

$$\begin{aligned}
L_{out}^{NOT}(x) &= -L_{in}(x),\\
L_{out}^{AND}(x_1,x_2) &= \max[L_{in}(x_1), L_{in}(x_2)],\\
L_{out}^{OR}(x_1,x_2) &= \min[L_{in}(x_1), L_{in}(x_2)],\\
L_{out}^{XOR}(x_1,x_2) &= sign[L_{in}(x_2)]\cdot sign[L_{in}(x_2)]\\
&\quad \cdot \min[L_{in}(x_1), L_{in}(x_2)],
\end{aligned} \tag{2}$$

where *max*() and *min*() function returns the higher and lower value of two variables, respectively, while the *sign*() function returns the sign of the variable.

B. Brief overview of the max-log MAP in LC method

In this method, the logic circuits create the output soft-value according to the following rules [8], [12]:

$$L_{out}^{\max-\log MAP\, in\, LC} = (R_1 - R_0)/2\,, \tag{3}$$

where the coefficients R_1 and R_0 are obtained, for *AND* logic circuits, as:

$$\begin{aligned}
R_1 &= L_{in}(\mathrm{x}_1) + L_{in}(\mathrm{x}_2),\\
R_0 &= \max[-L_{in}(\mathrm{x}_1) - L_{in}(\mathrm{x}_2), -L_{in}(\mathrm{x}_1) + L_{in}(\mathrm{x}_2),\\
&\qquad + L_{in}(\mathrm{x}_1) - L_{in}(\mathrm{x}_2)],
\end{aligned} \tag{4}$$

while for *OR* logic circuits as:

$$\begin{aligned}
R_1 &= \max[-L_{in}(\mathrm{x}_1) + L_{in}(\mathrm{x}_2), + L_{in}(\mathrm{x}_1) - L_{in}(\mathrm{x}_2),\\
&\qquad + L_{in}(\mathrm{x}_1) + L_{in}(\mathrm{x}_2)],\\
R_0 &= -L_{in}(\mathrm{x}_1) - L_{in}(\mathrm{x}_2),
\end{aligned} \tag{5}$$

and finally, for *XOR* logic circuits, using expressions:

$$\begin{aligned}
R_1 &= \max[-L_{in}(\mathrm{x}_1) + L_{in}(\mathrm{x}_2), + L_{in}(\mathrm{x}_1) - L_{in}(\mathrm{x}_2)],\\
R_0 &= \max[-L_{in}(\mathrm{x}_1) - L_{in}(\mathrm{x}_2), + L_{in}(\mathrm{x}_1) + L_{in}(\mathrm{x}_2)].
\end{aligned} \tag{6}$$

The *NOT* logic circuits work the same as for the *min-max in LC* method.

Several additional details on the *min-max in LC* and *max-log MAP in LC* methods can be found in several previously published papers [7]-[9].

C. The soft-decision MTR decoder

The soft-decision MTR decoder will be realized using the improved log circuits, with the implementation of one of soft-decision methods, the *min-max in LC* and *max-log MAP in LC* methods, as depicted in Figure 2.

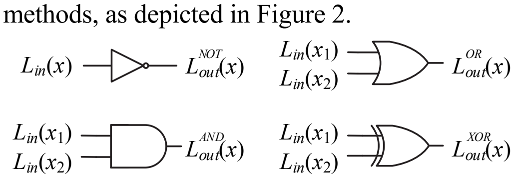

Figure 2. Redesigned Boolean logic circuits.

Using this approach, the idea is to preserve the simplicity of the hardware realization of the decoder and at the same time to design circuits capable of working with soft-values.

III. MTR ENCODING/DECODING OVER E^{2}PR4 CHANNEL

Performances of the *min-max in LC* and *max-log MAP in LC* were considered over the high density E^{2}PR4 magnetic recording channel, encoded by the simple LDPC-MTR serial concatenation. In this early stage of investigation, the simulation schemes do not employ any specially designed codes.

A. LDPC–MTR encoding process

The LDPC-MTR encoding was performed using R = 0.96 LDPC as an outer code and a well-known, rate 4/5 (2, k = 8) MTR, as an inner code, in the simulation scheme depicted in Figure 3.

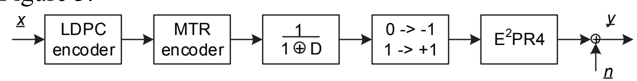

Figure 3. LDPC–MTR encoding over E^{2}PR4 channel.

The LDPC code of the length N = 4732, with M = 169 parity bits and column-weight 3 is based on Kirkman triple systems [13].

Unfortunately, the resulting code rate of concatenation is only R = 0.771. This is not quite suitable for high density magnetic recording; however, it is not limiting for such initial investigation of MTR soft-decision decoding methods.

It was assumed that the read-back signal y is distorted with the additive white Gaussian (AWGN) noise n and that the signal-to-noise ratio (SNR) is defined as

$$SNR = 10\log(E_b / N_o) = 10\log(E_c / 2R\sigma^2)\,, \tag{7}$$

where $E_c = RE_b$ is the symbol bit energy at channel output, N_o is one-sided power spectral density and σ^2 is noise variance.

B. LDPC–MTR decoding process

The channel detection was done using the Soft-Output Viterbi Algorithm (SOVA), with the detection window of 20 symbols, while the LDPC code was decoded by the message-passing algorithm [14].

During the decoding process, the soft information was exchanged between SOVA decoder and message-passing, as depicted in Figure 4.

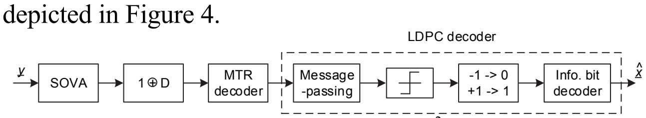

Figure 4. LDPC–MTR decoding over E^{2}PR4 channel.

Intention of this paper was to investigate the performance of the proposed methods for the MTR soft-decision decoding. Thus, the simulation scheme implements the reverse and straightforward concatenation of LDPC-MTR, without any iterative exchange between them, focusing on the ability of soft-values propagation through the MTR decoder.

IV. THE COMPLEXITY OF THE PROPOSED METHODS

Both methods suggest the redesign and enhancement of the conventional logic circuits, though they determine the output soft-values for particular logic circuits in a different way. This section is intended to emphasize the difference between them, as well as their complexity.

A. AND logic circuits

The analysis of difference between the output soft-values of methods, for *AND* circuits, was performed considering that both input soft-values range between -8 and +8, with the step of 0.1. The major difference can be noticed when inputs have the same value but different sign, as depicted in Figure 5, part *a*.

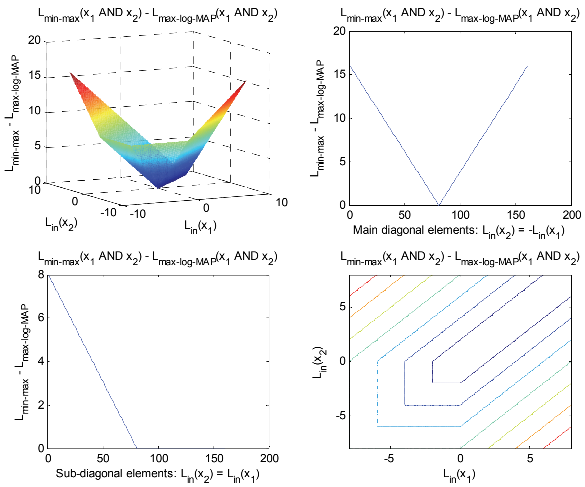

Figure 5. Difference of the output soft-value for AND circuits.

Moreover, analyzing the two-dimensional matrix, whose columns and rows represent $L_{in}(x_1)$ and $L_{in}(x_2)$, it can be noticed that over the main diagonal, where $L_{in}(x_2) = -L_{in}(x_1)$, the difference between methods is very high and symmetrically distributed over the point where $L_{in}(x_1)$ and $L_{in}(x_2)$ are equal to zero, as depicted in Figure 5, part *b*.

Furthermore, over the sub-diagonal of the matrix, where $L_{in}(x_2) = +L_{in}(x_1)$, the difference between methods is lower and exists only when input values are negative, as depicted in Figure 5, part *c*. In case that both input soft-values are positive, the *min-max in LC* and *max-log MAP in LC* identically assess the output soft-values.

Analyzing other points in $[L_{in}(x_1), L_{in}(x_2)]$ matrix, which are off-diagonal elements, it can be noticed that certain positive differences exist, as depicted in Figure 5, part *d*. This difference is not extreme as those on the main diagonal and thus it can be concluded that both of these two methods asses correctly the output soft-value, but with different confidence.

Even both proposed methods use the *min*() and *max*() functions, it seems that the *min-max in LC* approach produces some higher soft-values on the output of the *AND* logic circuit than the *max-log MAP in LC* method. Only in a case when both input values $L_{in}(x_1)$, $L_{in}(x_2)$ are equal and positive, these two methods produce the same output soft-value.

B. OR logic circuits

The similar analysis of difference between assessment in the suggested methods can be performed for *OR* logic circuits, as depicted in Figure 6.

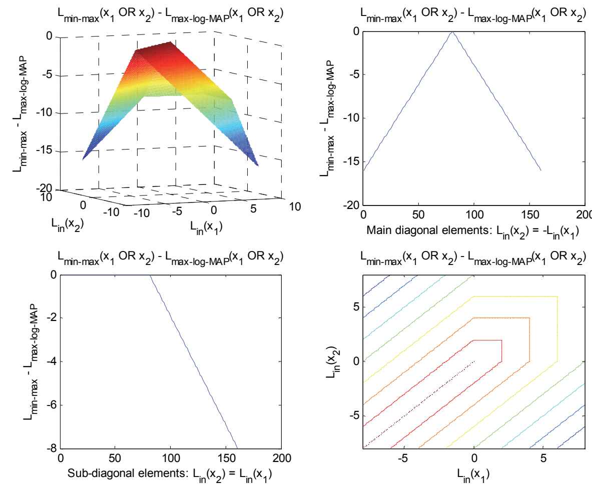

Figure 6. Difference of the output soft-value for OR circuits.

In this case, the *min-max in LC* assesses lower output soft-values, compared to *max-log MAP in LC*, with the similar situation that the major difference lies on the main diagonal of $[L_{in}(x_1), L_{in}(x_2)]$ matrix, where $L_{in}(x_2) = -L_{in}(x_1)$. Also, we have symmetric distribution over the point where $L_{in}(x_1)$ and $L_{in}(x_2)$ are equal to zero, as depicted in Figure 6, part *b*.

Additionally, on the sub-diagonal of $[L_{in}(x_1), L_{in}(x_2)]$ matrix, both methods equally assess output values, when input values are equal and negative, as depicted in Figure 6, part *c*.

Moreover, analyzing the shape of the difference function, it can be concluded, again, that both methods similarly assess soft-values on the circuit output, unfortunately, with quite different confidence when $L_{in}(x_2) = -L_{in}(x_1)$.

C. XOR logic circuits

In the case of *XOR* logic circuits, the *min-max in LC* and the *max-log MAP in LC* methods assess output soft-values with almost negligible difference, which ranges around 10^{-15}, as depicted in Figure 7, part *a*. Thus, both the proposed methods are equally confident and offer a quite suitable soft-value for *XOR* circuits.

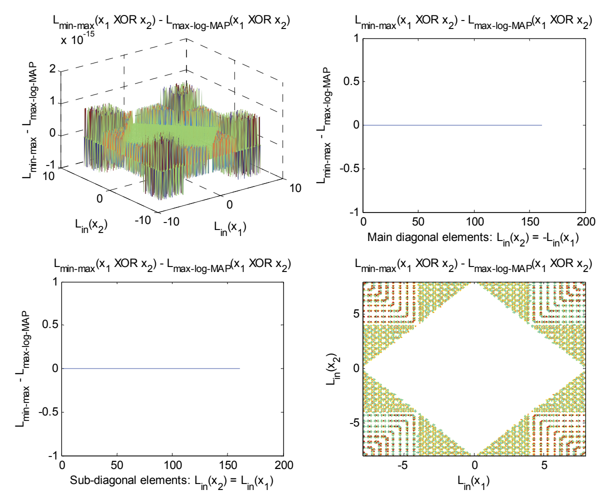

Figure 7. Difference of the output soft-value for XOR circuits.

Even more, there is no any difference on the main and sub-diagonal of $[L_{in}(x_1), L_{in}(x_2)]$ matrix, as depicted in Figure 7, parts *b* and *c*, while on a wide range of the off-diagonal elements this difference is zero, as shown in Figure 7, part *d*.

These two proposed methods have suitable assessments of the soft-values on the circuit outputs. However, during their implementation into the new MTR decoder, the attention has to be paid to the complexity and the total number of the required operations in the decoder.

D. *Analysis on the required number of operations*

The *min-max in LC* and *max-log MAP in LC* methods are based on the *min*() and *max*() functions in order to assess the soft-values. Regarding the number of operations, the *min-max in LC* is in slight advantage, as depicted in Figure 8.

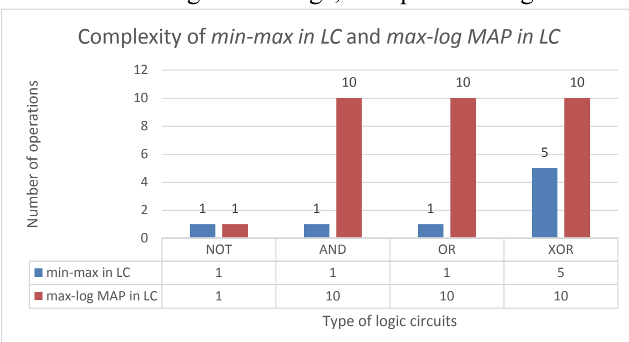

Figure 8. Required number of operations.

Unfortunately, the several required operations with input soft-values, in the *max-log MAP in LC*, increase the number of overall operations, even though the complexity of both methods is still of the same degree. However, regarding the assembly of the new MTR decoder that is based on the soft-decision approach, the total number of the required operations by both proposed methods in all logic circuits should be considered, since that can have a substantial impact on the power consumption of the newly created decoder.

Analyzing the decoder for the well know rate 4/5 (2, k = 8) MTR code, depicted in Figure 1, it can be noticed that such realization requires 5 *NOT*, 8 *AND* and 5 *OR* logic circuits, with a total number of decoder operations as depicted in Figure 9.

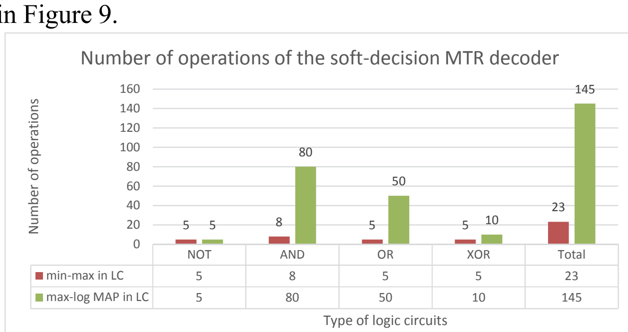

Figure 9. Total number of operations for the MTR decoder.

Additionally, the *min-max in LC* and *max-log MAP in LC* methods require one *XOR* circuit in the data sequence pre-processing phase, before the sequence enters into the MTR decoder. The *XOR* circuit is not depicted in the MTR decoder, presented in Figure 1, though it should be counted in the total number of the MTR decoding process.

Even though these proposed soft-decision methods have the same degree of complexity, the *min-max in LC* approach requires 6.3 times less operations than the other one. Therefore, the *min-max in LC* outperforms the *max-log MAP in LC* and will be more desirable for the practical implementation into the soft-decision MTR decoder.

V. SIMULATION RESULTS

In order to provide the final conclusion on *min-max in LC* and *max-log MAP in LC* approach for soft-decision decoding of the MTR codes, the comparison was made over the one-track E^2PR4 magnetic recording channel, utilizing serial concatenation of LDPC-MTR codes. Performances of these methods have been summarized in Figure 10.

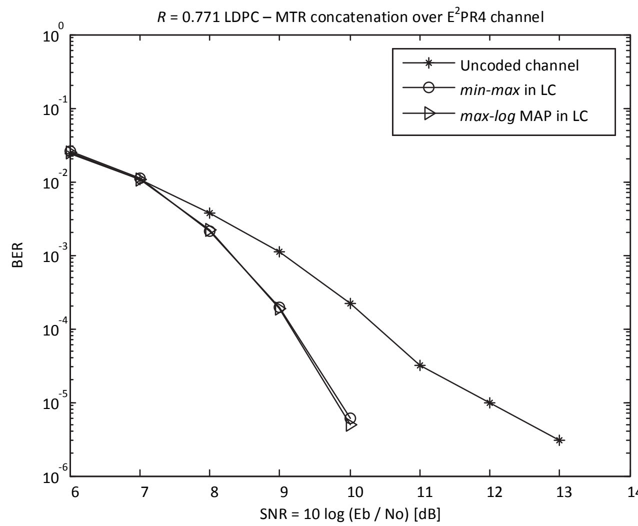

Figure 10. Approaches for MTR soft-decision decoding.

It can be observed that the utilization of both *min-max in LC* and *max-log MAP in LC* approaches in the logic circuits of the new MTR decoder, within the LDPC-MTR serial concatenation, resulted with almost identical and, in addition, the considerable coding gain of 2.2 dB for BER = 10^{-5}, comparing with the uncoded channel.

This result confirms that both methods properly estimate soft-values on the logic circuits outputs, allowing the propagation of the soft-information to the LDPC decoder.

Using both these approaches, the new MTR decoder can be equipped with the soft-decision decoding capability, offering the confidence of binary decisions per output, which is required for the iterative simulation schemes.

Both methods are almost identical with the achieved decoding gain, but a slight advantage is on the side of the *min-max in LC*. This approach resulted with 6.3 time lower number of operations in the MTR decoder, which can be quite valuable in case of extensive utilization of decoder and demanding signal processing. The lower number of arithmetic operations can have strong influence on the power consumption of practical MTR decoder.

 ISBN: 978-1-61208-402-2

VI. Conclusion

This paper presents a comparison of two newly proposed methods, *min-max in LC* and *max-log MAP in LC* methods, for the soft-decision decoding of the MTR code. Both methods suggest the redesign and enhancement of the classical Boolean logic circuits and their implementation into the new MTR decoder.

Both approaches achieved almost identical decoding gain in a serial concatenation of the MTR and LDPC codes, over the E^2PR4 magnetic recording channel. The obtained decoding gain confirms that proposed methods appropriately evaluate soft-values on the outputs of the logic circuits in a soft-decision MTR decoder, enabling their implementation in an iterative decoding scheme.

However, the slight advantage is on the side of the *min-max in LC* method, regarding the required number of operations in a new MTR decoder. The reason for this lies in the fact that the simple *min*() and *max*() decisions are required in the *min-max in LC* method in order to assess the soft-value on the output of logic circuits, compared to the *max-log MAP in LC*, which requires additional operations with the input soft-values before the decision is made.

The low number of operations recommends the *min-max in LC* method as a fairly suitable for the implementation into the soft-decision MTR decoder, which potentially can have relatively small power consumption.

This paper has indicated also, that *min-max in LC* method is about to produce the less confident output soft-values, for *AND* and *OR* logic circuits, when the input soft-values are equal, though with a different sign, $L_{in}(x_2) = -L_{in}(x_1)$. Thus, this issue can be a foundation for future research on the *min-max in LC* enhancement.

Acknowledgment

This paper has been supported by the Provincial Secretariat for Science and Technological Development of the Autonomous Province of Vojvodina, the Republic of Serbia, through the grant for the project 114-451-2061/2011-01.

References

[1] J. Moon and B. Brickner, “Maximum transition run codes for data storage systems,” IEEE Trans. Magn., vol. 32, no. 5, pp. 3992-3992, September 1996.

[2] D. J. C. MacKay and R. Neal, “Near Shannon limit performance of low density parity check codes,” IEE Electron. Letters, vol. 33, pp. 457-458, March 1997.

[3] R. M. Todd and R. Cruz, "Enforcing maximum-transition-run code constraints and low-density parity-check decoding," IEEE Trans. Magn., vol. 40, no. 6, pp. 3566-3571, November 2004.

[4] T. Kanaoka and T. Morita, "Structured LDPC Codes With Reversed MTR/ECC for Magnetic Recording Channels," IEEE Trans. Magn., vol. 42, no. 10, pp. 2561 - 2563, October 2006.

[5] M. Blaum, R. Galbraith, K. Lakovic and B. Wilson, "Reverse Concatenation with Maximum Transition Run (MTR) Codes for High-Density Perpendicular Recording," IEEE Global Telecommunications Conference – GLOBECOM 2008, November 30 – December 4, 2008, New Orleans, USA, pp. 1-4.

[6] J. Hagenauer, “Source-controlled channel decoding,” IEEE Trans. Comm., vol. 43, No. 9, pp. 2449-2457, September 1995.

[7] N. Djuric and M. Despotovic, “Soft-output decoding in multiple-head MTR encoded magnetic recording Systems,” IEEE International Conference on Communications – ICC 2006, Istanbul, Jun 11 – 15, 2006, vol. 3, pp. 1255-1258.

[8] N. Djuric, “MAP decoding of MTR codes in LDPC-MTR encoded magnetic recording systems,” in Proc. 9th International Conference on Telecommunications in Modern Satellite, Cable and Broadcasting Services – TELSIKS 2009, Nis, Serbia, 7-9 October, 2009, vol. 34, no. 3, pp. 381- 392.

[9] N. Djuric and V. Senk, “The MAP in LC Decoding of MTR Codes in Two-Track Magnetic Recording Systems”, The 6th International Conference on Communication Theory, Reliability, and Quality of Service – CTRQ 2013, April 21-26, 2013, Venice, Italy, pp. 58-63.

[10] H. K. Thapar and A. M. Patel, “A class of partial-response systems for increasing storage density in magnetic recording”, IEEE Trans. Magn., vol. MAG-25, pp. 3666-3668, Sep. 1987.

[11] N. Djuric and M. Despotovic, “Application of MTR soft-decision decoding in multiple-head magnetic recording systems,” Indian Academy of Sciences, Sadhana – Academy Proceedings in Engineering Science, vol. 34, Part 3, pp. 381–392, June 2009.

[12] N. Djuric and V. Senk, “The MAP in LC As a Soft-decision Method for MTR Decoding in LDPC-MTR Concatenation”, World Symposium On Computer Networks and Information Security International Conference on Computer and Computational Intelligence – ICCCI 2014, Hammamet, Tunisia, June 13-15, 2014, vol. 1, pp. 28, 1-4.

[13] B. Vasic, E. Kurtas and A. Kuznetsov, "Kirkman systems and their application in perpendicular magnetic recording," IEEE Trans. Mag., vol. 38, no. 4, Part: 1, pp. 1705 -1710, July 2002.

[14] F. R. Kschischang, B. J. Frey and H.-A. Loeliger, “Factor graphs and the sum-product algorithm,” IEEE Trans. Inform. Theory, pp. 498-519, February 2001.

 ISBN: 978-1-61208-402-2