Universität Stuttgart

INSTITUT FÜR
KOMMUNIKATIONSNETZE
UND RECHNERSYSTEME
Prof. Dr.-Ing. Andreas Kirstädter

Institute of Communication Networks and Computer Engineering
University of Stuttgart
Pfaffenwaldring 47, D-70569 Stuttgart, Germany
Phone: ++49-711-685-68026, Fax: ++49-711-685-67983
Email: mail@ikr.uni-stuttgart.de, http://www.ikr.uni-stuttgart.de

Reducing Network Operators' Expenses by Adjusting the MTTR

Tobias Enderle
Institute of Communication Networks
and Computer Engineering (IKR)
University of Stuttgart
Stuttgart, Germany
tobias.enderle@ikr.uni-stuttgart.de

Andreas Kirstädter
Institute of Communication Networks
and Computer Engineering (IKR)
University of Stuttgart
Stuttgart, Germany
andreas.kirstaedter@ikr.uni-stuttgart.de

***Abstract*—Network operators have to generate sustainable profits from the sale and provision of network connections. However, due to quickly growing bandwidth requirements and stagnating revenues, this becomes increasingly difficult to achieve. Lowering expenses is a proven remedy in such situations. The major part of a network operator's expenses is related to the provision of network connections and, in particular, their quality of service (QoS) in terms of availability. Here, an important aspect is the repair of network components to restore connectivity in case of failures. Lowering expenses for network repair clearly results in a reduction of total enterprise expenses. However, it also results in an increase in compensation payments because availability guarantees will be violated more frequently. This paper explores the trade-off between the reduction in repair-related expenses and the corresponding increase in compensations. To this end, a general financial model is developed which considers the described trade-off. The model involves corporate key figures, a cost function that associates repair-related expenses with the mean time to repair (MTTR), and a probabilistic estimation of the expected compensations. Using exemplary model parameters, we show that the increase in compensations typically does not outweigh the reductions in repair-related expenses. Consequently, significant expense reductions are possible for network operators.**

***Index Terms*—Availability management, Compensation, Cost reductions, Mean time to repair (MTTR), Profit margin, Service level agreement (SLA)**

I. INTRODUCTION

The main goal of every commercial enterprise is to achieve sustainable profit. This can be reached by solid revenues and permanent monitoring of the expenses. Due to rapidly increasing bandwidth requirements and stagnating revenues, especially network operators have to observe their operational and capital expenses intensively. An important aspect in this field is the service quality offered in terms of end-to-end path availability. Addressed by service level agreements (SLAs), the operator guarantees a specific availability measured via the outage durations per billing cycle experienced by the customer. The availability levels achieved are directly influenced by the mean time to repair (MTTR) values that the operator is targeting. Clearly, the involvement of more repair staff or the usage of higher quality line and component installations will reduce the MTTR and increase the availability levels that can be achieved and guaranteed. Accordingly, this increased service level in terms of higher availability is related to higher production expenses in terms of salary and equipment. On the other hand, lowered availability levels will increase the compensation that the operator has to pay if it cannot fulfill its SLAs.

Economic strategies for highly survivable networks are of great interest. So, e.g., in [1] the optimum allocation split of additional investment between MTTR reduction on one side and increased double-failure restorability on the other side was investigated. An analytical framework was established for networks that already show $100\,\%$ single failure restorability. It was shown that it is preferable to give more invest to MTTR reduction than to the improvement of double failure restorability. The authors of [2] extended the multi-period planning of network capacity upgrades to analyze the trade-off between the savings gained from postponed investment in protection upgrades and the resulting SLA compensations. In this way, the optimum point in time for protection investment depending on the individual parameters (like cost decreasing factor, failure rate, agreed service availability etc.) could be identified. The influence of the outage distribution on outage-related expenses was investigated in [3] for the neighboring discipline of IT systems. Via a detailed framework on IT service availability, it was shown that at the same availability level shorter outages are to be preferred.

This paper will consider closely the influence of the targeted MTTR level on the trade-off between production expenses and SLA compensation on a general level. We will demonstrate via a ceteris paribus analysis that the benefit from a reduced MTTR performance in terms of reduced production expenses (for network services) will strongly outweigh the increase of SLA compensation that has to be paid. For this purpose, we will develop a general financial model in Section II and consider the influence of the MTTR level on production expenses in Section III. In Section IV we develop from the distributions of the outage durations a probabilistic model for the compensations depending on the MTTR. Both expense components are then evaluated jointly in Section V.

This work has been performed in the framework of the CELTIC EUREKA project AI-NET-ANTILLAS (Project ID C2019/3-3), and it is partly funded by the German BMBF (Project ID 16KIS1312). The authors alone are responsible for the content of the paper.

II. GENERAL FINANCIAL MODEL

To be able to analyze the trade-off between production cost and SLA compensations we need a suitable model. Starting at the enterprise level, the enterprise profit margin M is the company key figure. It is calculated as the return B before tax and financing charges normalized to the revenue V. In the case of network operators, we find values for M on the order of a few percent (e.g. [4], [5])—typical for companies in mature markets. The return B itself is calculated by subtracting from V the sum of the different expense positions E of the company:

$$M = \frac{B}{V} = \frac{V - E}{V}. \tag{1}$$

E can be differentiated into expenses for administration (E_A), sales (E_S) and the production-related expenses E_P for the output of the company, i.e., its goods and services [6]. Generally, at network operators—like other production enterprises—E_P will be by far the largest part. At a closer look we see that E_P plus some parts of E_A and E_S will be directly influenced by the chosen level of MTTR: More spare hardware, higher quality and serviceability of production equipment (i.e., installed network hardware and software) will lead both to higher expenses and MTTR reduction. This effect is especially visible for service staff—more staff closer to the network installations creates higher expenses but clearly speeds the reaction to outages. These effects even get visible for E_A and E_S. Here, more or better trained staff will also contribute partly to MTTR reduction. In the next section, we will present a model to relate the sum of all MTTR-related expenses, which we denote by E_R in the following, to the MTTR level. At this moment, we can state that E_R will amount to a large share of E—far more than the cost of the actual repair activities alone. Due to the small values of the profit margin M mentioned above, E and V are very close. Thus, we can say that E_R will also amount to a very large part of V.

Reviewing the compensation due to SLA violations offers another interesting insight: Compensation is usually paid as a fraction of the periodical revenues (per billing cycle) per customer service. The exact value of this depends on the extent of violation of the agreed availability experienced by the customer (see Section IV below). This compensation directly reduces V. A positive profit margin, M, can only result if the revenue, V, is larger than the total expenses, E. If we consider the small levels of M mentioned above, we directly see that already a small (e.g. $< 1\,\%$) reduction of V would have detrimental consequences for the overall profitability of the enterprise. Therefore, the compensation expenses E_C must stay limited to a very small fraction of E and V. Considering the fact that E_R is the dominant share of E and V, the compensation expenses E_C will also be much smaller than E_R—offering some playground for optimization that we will explore in Section V.

Following a ceteris paribus analysis, we consider the company at a fixed "operation point" in terms of V. This means that an increase in MTTR and the resulting decrease in E_R will not change the overall value of V—especially any additional customer churn is assumed to be negligible. We denote with the variables E_C and $E_{\mathrm{C},0}$ the compensations and their original value before the MTTR increase. Similarly, we denote with E_R and $E_{\mathrm{R},0}$ the MTTR-related expenses and their original value before the MTTR increase. We define their relative changes as:

$$q_\mathrm{C} = \frac{E_\mathrm{C}}{E_{\mathrm{C},0}} - 1 \tag{2}$$

$$\text{and} \quad q_\mathrm{R} = 1 - \frac{E_\mathrm{R}}{E_{\mathrm{R},0}}. \tag{3}$$

With E_T we denote the sum of E_C and E_R:

$$E_\mathrm{T} = E_\mathrm{C} + E_\mathrm{R}. \tag{4}$$

With these definitions and the considerations from above, we can specify the relation between both expense positions $E_{\mathrm{C},0}$ and $E_{\mathrm{R},0}$ as the parameter ψ:

$$\psi = \frac{\frac{E_{\mathrm{C},0}}{V}}{\frac{E_{\mathrm{R},0}}{V}} = \frac{E_{\mathrm{C},0}}{E_{\mathrm{R},0}}. \tag{5}$$

The relative change of E_T:

$$q_\mathrm{T} = 1 - \frac{E_\mathrm{T}}{E_{\mathrm{T},0}} \tag{6}$$

can then be written as:

$$\begin{aligned} q_\mathrm{T} &= 1 - \frac{E_\mathrm{C} + E_\mathrm{R}}{E_{\mathrm{C},0} + E_{\mathrm{R},0}} & (7)\\ &= 1 - \frac{E_{\mathrm{C},0} \cdot (q_\mathrm{C} + 1) + E_{\mathrm{R},0} \cdot (1 - q_\mathrm{R})}{E_{\mathrm{C},0} + E_{\mathrm{R},0}} & (8)\\ &= 1 - \frac{\psi E_{\mathrm{R},0} \cdot (q_\mathrm{C} + 1) + E_{\mathrm{R},0} \cdot (1 - q_\mathrm{R})}{\psi E_{\mathrm{R},0} + E_{\mathrm{R},0}} & (9)\\ &= \frac{q_\mathrm{R} - \psi q_\mathrm{C}}{\psi + 1}. & (10) \end{aligned}$$

A positive value of q_T marks a reduction in the combined expenses E_T. This is what a network operator wants to achieve. Since $\psi \geq 0$, we get $q_\mathrm{T} > 0$ if $q_\mathrm{R} > \psi \cdot q_\mathrm{C}$. The relative reduction in MTTR-related expenses cannot exceed $100\,\%$. Consequently, $q_\mathrm{R} \leq 1$. On the other hand, the relative increase in compensations, q_C, can be much higher than $100\,\%$. In the worst case, compensations can eat up all revenues V. Therefore, it is crucial to evaluate and compare the behavior of q_R and q_C in detail. In the next sections, we extend our model to allow such a comparison.

III. REDUCTION OF MTTR-RELATED EXPENSES

The MTTR-related expenses, E_R, make up a large share of the total expenses a network operator has to face. Therefore, the reduction of those expenses is a promising way to increase the profit margin. In order to estimate the relationship between reductions in MTTR-related expenses and the resulting MTTR in the network, we need a suitable model. Following the argumentation of Grover [1], we assume that the expenses to provide an infinite MTTR vanish because no repair is needed at all. At the other extreme, we assume that there exists a minimum MTTR even with infinite costs. There is no way to reduce the MTTR below this minimum because not only would

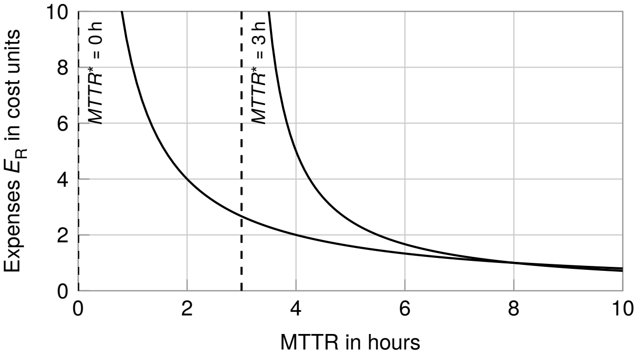


Fig. 1. Expenses E_R to achieve a certain level of MTTR.

this require a huge amount of spare hardware but also repair personnel at every possible failure location let alone instant repair methods. In between these two extremes, we assume that further MTTR reductions become increasingly costly the smaller the MTTR already is. Therefore, we suggest a model for the MTTR-related expenses based on a hyperbolic function:

$$E_R(MTTR) = \frac{\alpha}{(MTTR - MTTR^*)} \tag{11}$$

where $MTTR^*$ is the minimum attainable MTTR and α is a scaling parameter. Figure 1 shows two exemplary cost functions with $MTTR^* = 0\,\text{h}$ and $3\,\text{h}$, respectively. For this example, the parameter α has been chosen such that $E_R(8\,\text{h}) = 1$ cost unit.

If the network operator lowers the MTTR-related expenses, the MTTR for its network increases. According to (3) and with the cost function in (11) the relative reduction in MTTR-related expenses for an MTTR increase by a factor of $f > 1$ is given by

$$\begin{aligned} q_R &= 1 - \frac{E_R(f \cdot MTTR_0)}{E_R(MTTR_0)} && (12) \\ &= 1 - \frac{\frac{\alpha}{(f \cdot MTTR_0 - MTTR^*)}}{\frac{\alpha}{(MTTR_0 - MTTR^*)}} && (13) \\ &= 1 - \frac{1 - \frac{MTTR^*}{MTTR_0}}{f - \frac{MTTR^*}{MTTR_0}} && (14) \\ &= 1 - \frac{1 - \Phi}{f - \Phi} && (15) \end{aligned}$$

where $MTTR_0$ is the initial MTTR in the network, $MTTR^*$ is the minimum attainable MTTR and $\Phi = MTTR^*/MTTR_0$. Since q_R is independent of the parameter α, the only parameter that needs to be selected for the cost model is $MTTR^*$. With increasing MTTR, the MTTR-related expenses shrink. Therefore, the quantity q_R is non-negative.

Figure 2 shows the relative expense reductions, q_R, over the relative MTTR increase, f, for different values of Φ. The expense reductions grow both with the MTTR increase, f, and with the ratio of initial to minimum attainable MTTR, Φ. The higher Φ is, the closer the initial MTTR, $MTTR_0$, is to the minimum attainable MTTR, $MTTR^*$. The cost function (Figure 1) is steep in those regions close to $MTTR^*$, which means that even small MTTR increases lead to large expense reductions. We assume that a network operator is able to estimate Φ for its own network and services. In general, however, the smaller Φ is, the smaller the attainable expense reductions are. Therefore, $\Phi = 0$ represents the most conservative assumption and it prevents overestimation of the expense reductions. Since $\Phi = 0$ implies $MTTR^* = 0\,\text{h}$, it is also in line the assumptions made in [1].

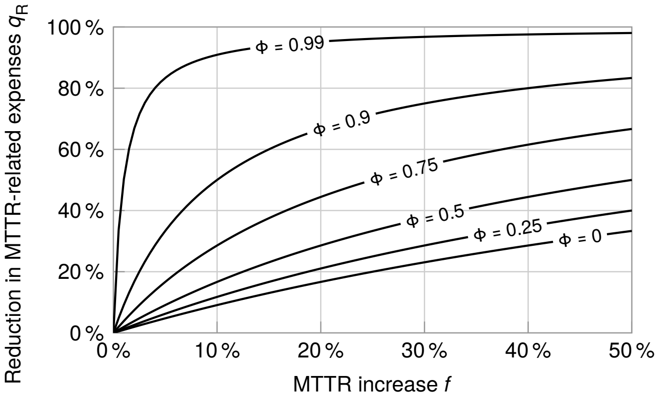


Fig. 2. Reduction in MTTR-related expenses on MTTR increase.

IV. ESTIMATION OF EXPECTED COMPENSATIONS

Network services are composed of many different components, such as fibers, network devices, and the corresponding software. These components are subject to failures and service downtime is inevitable, even though network services are often protected by redundant backup components. If a customer's network service experiences downtime, the network operator has to pay compensations in the form of refunds or service credits to this customer. The relationship between downtime and compensation amount is defined in the SLA. Usually, the compensation is determined at the end of a billing cycle as a function of the cumulated downtime (or unavailability, respectively) during that cycle. The length of a billing cycle is typically one calendar month and the compensation amount is a percentage of the monthly recurring charge (MRC) the customer pays for the service. Figure 3 depicts a policy for a protected and a policy for an unprotected network service. The two policies are derived from actual policies for wavelength services found in SLAs of large network operators [7]–[9]. The step function design is typical for these policies.

Since network components fail randomly, we model the cumulated downtime per billing cycle as the random variable (RV) X and the corresponding compensations as RV C. The compensation policy, i.e., the mapping from X to C, can be defined as $g : [0, T] \to \mathbb{R}$ with

$$g(x) = \begin{cases} 0 & \text{for } x \le x_1 \\ c_i & \text{for } x_i < x \le x_{i+1}, \quad 1 \le i < n \\ c_n & \text{for } x > x_n \end{cases} \tag{16}$$

where $T = 1$ month, x_i are the thresholds of monthly downtime in the SLA and c_i are the corresponding compensation

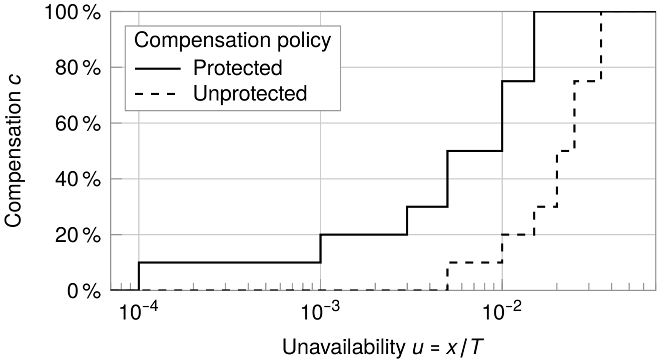


Fig. 3. Generic compensation policies for protected and unprotected services. The y-axis is given as a percentage of the monthly recurring charge (MRC). The x-axis is in log scale.

TABLE I
GENERIC COMPENSATION POLICIES FOR PROTECTED AND UNPROTECTED SERVICES AND A BILLING CYCLE OF ONE MONTH.

Policy	Availability	Unavailability	Downtime per month (x_i)	Compensation as percentage of MRC (c_i)
Protected	0.9999	0.0001	0.073 h	10 %
	0.999	0.001	0.73 h	20 %
	0.997	0.003	2.19 h	30 %
	0.995	0.005	3.65 h	50 %
	0.99	0.01	7.3 h	75 %
	0.985	0.015	10.95 h	100 %
Unprotected	0.995	0.005	3.65 h	10 %
	0.99	0.01	7.3 h	20 %
	0.985	0.015	10.95 h	30 %
	0.98	0.02	14.6 h	50 %
	0.975	0.025	18.25 h	75 %
	0.965	0.035	25.55 h	100 %

levels. The parameter n denotes the number of downtime thresholds in the policy ($n = 6$ in Figure 3). Table I shows the values that define the policies of Figure 3 in more detail, in particular the values for x_i and c_i used in (16).

In order to compute the expected compensation for a service we need to know the distribution of the cumulated downtime X. To this end, we assume that the network service can be in one of two states: working or not working. The time per month the service is not working constitutes the cumulated downtime. We assume that the time before a failure and the time it takes to repair the service both follow exponential distributions and the mean time to failure (MTTF) and MTTR are their respective expected values. In [10], Takács considers the cumulated downtime Z in a time interval D for a system with failure rate γ and repair rate δ. Under the condition that this system is in the working state at the beginning of the time interval D, the cumulative distribution function (CDF) of Z is given by [10]

$$\Omega_{\gamma,\delta,D}(z) = e^{-\gamma(D-z)}\left(1+\sqrt{\gamma\delta(D-z)} \cdot \int_0^z e^{-\delta y} y^{-\frac{1}{2}} I_1\left(2\sqrt{\gamma\delta(D-z)y}\right) \mathrm{d}y\right) \tag{17}$$

where $I_1(x)$ is the modified Bessel function of the first kind of order 1. The assumption that the service is working at the beginning of the time interval D is valid if we consider a service model with a single billing at the end of the contract period and we consider this contract period to be the interval D. In this model, the service is working at the beginning when it is handed over to the customer. Billing and compensation are done once at the end of the contract period when the service leaves the network.

In this work, however, we consider the case that the customer is billed every month, i.e., $D = T$, and the service remains in the network for multiple billing cycles, e.g., with a contract period of 1, 2 or 5 years. Consequently, a service can be down at the beginning of a billing cycle, namely when a failure has not been repaired until the end of the previous billing cycle. Due to this difference, the cumulated downtime per billing cycle, X, cannot be modeled by the cumulated interval downtime, Z. Instead, the distribution of X has been derived in [11] as an extension to the distribution of Z. Following [11], the CDF of X is given by

$$F(x) = a\Omega_{\lambda,\mu,T}(x) + (1-a)\big(1-\Omega_{\mu,\lambda,T}(T-x)\big) \tag{18}$$

where $a = \mathit{MTTF}/(\mathit{MTTF} + \mathit{MTTR})$ is the steady-state availability of the service, $\lambda = 1/\mathit{MTTF}$, and $\mu = 1/\mathit{MTTR}$. The billing cycle T is 1 month in our case.

With the policy function, $g(x)$, and the CDF of the cumulated downtime, $F(x)$, the expected value of the compensation $C = g(X)$ can be calculated by piecewise integration as

$$\mathbb{E}[C] = \int_0^T g(x)\,\mathrm{d}F(x) \tag{19}$$

$$= 0 + \sum_{i=2}^{n} \int_{x_{i-1}}^{x_i} g(x)\,\mathrm{d}F(x) + \int_{x_n}^{T} g(x)\,\mathrm{d}F(x) \tag{20}$$

$$= \sum_{i=2}^{n} c_{i-1}\big(F(x_i) - F(x_{i-1})\big) + c_n\big(F(T) - F(x_n)\big) \tag{21}$$

$$= \sum_{i=2}^{n} (c_{i-1} - c_i)F(x_i) - c_1 F(x_1) + c_n. \tag{22}$$

$\mathbb{E}[C]$ can be computed precisely, e.g., using the open-source Python package *SciPy* [12].

From (18) we can see that a change in the MTTR directly impacts $F(x)$ and, consequently, also $\mathbb{E}[C]$. In (2) we defined the relative increase in compensation as $q_\mathrm{C} = E_\mathrm{C}/E_{\mathrm{C},0} - 1$. Now, we set $E_\mathrm{C} = \mathbb{E}[C]$ and $E_{\mathrm{C},0} = \mathbb{E}[C_0]$ where C_0 is the RV representing the compensation amount for the initial MTTR_0 and C is the RV representing the compensation amount for $f \cdot \mathit{MTTR}_0$, the increased MTTR. Hence, for the relative increase in compensation we have

$$q_\mathrm{C} = \frac{\mathbb{E}[C]}{\mathbb{E}[C_0]} - 1. \tag{23}$$

With increasing MTTR, the expected amount of compensation increases as well. Therefore, the quantity q_C is non-negative.

TABLE II
REQUIRED MTTF FOR DIFFERENT INITIAL MTTR VALUES AND $\mathbb{E}[C_0] = 0.001 \cdot \mathit{MRC}$.

Policy	MTTR_0	MTTF	λ	Availability
Protected	2 h	174 566 h	5728 FIT	0.999989
	5 h	316 018 h	3164 FIT	0.999984
	12 h	487 190 h	2053 FIT	0.999975
	24 h	600 885 h	1664 FIT	0.99996
Unprotected	2 h	14 463 h	69 140 FIT	0.999862
	5 h	73 938 h	13 525 FIT	0.999932
	12 h	227 742 h	4391 FIT	0.999947
	24 h	395 738 h	2527 FIT	0.999939

FIT: Failures in time

V. ILLUSTRATIVE RESULTS

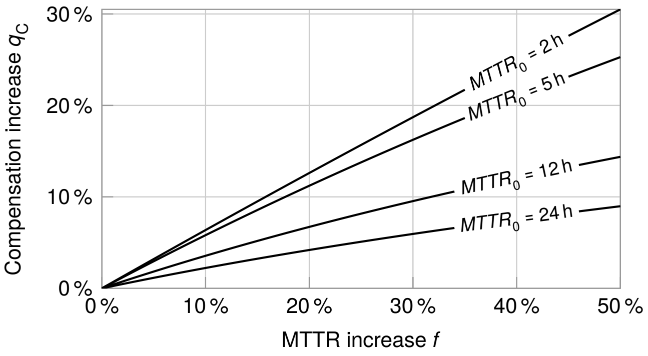
In the previous sections we have introduced the models for the reduction of MTTR-related expenses and the increase in compensations. Using these models, we now turn to illustrative numerical examples of the resulting expense reductions. We will first consider the behavior of the increase in compensations and put it into perspective from a qualitative point of view. After that, we evaluate the attainable reductions in combined expenses, q_T, numerically and provide detailed insights.

First, we consider the increase in compensations, q_C. The expected amount of compensation $\mathbb{E}[C]$ depends on the compensation policy as well as the MTTF and MTTR of the network service. We assume that the MTTF remains unchanged when the MTTR is increased from a value of MTTR_0 to a value of $f \cdot \mathit{MTTR}_0$. For MTTR_0, we consider the values 2 h, 5 h, 12 h and 24 h [13]. For each of those values we set the MTTF such that $\mathbb{E}[C_0] = 0.001 \cdot \mathit{MRC}$, i.e., we set it such that the initial compensations amount to $0.1\,\%$ of the total revenue V. The different MTTF values, their corresponding failure rate, and also the resulting steady-state availability are provided in Table II.

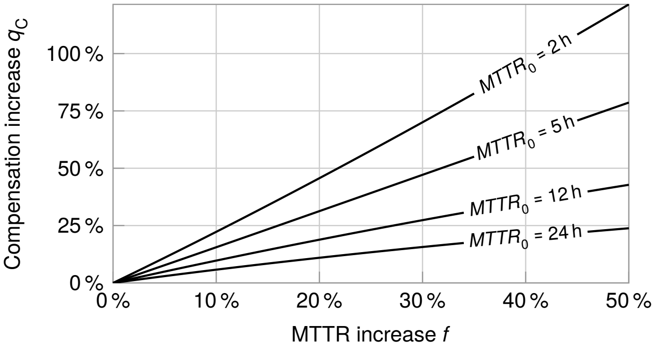
Figures 4a and 4b show the resulting increase in expected compensation, q_C, for the generic protected and unprotected compensation policies introduced in Section IV. It can be seen that the compensation increase is steeper for smaller values of MTTR_0. Also, for the policy of protected services (Figure 4a), the compensation increase is less sensitive to MTTR increases than for the policy of unprotected services (Figure 4b).

In Section II it was discussed that the increase in compensations, q_C, is virtually unlimited, while the reductions in MTTR-related expenses, q_R, cannot exceed $100\,\%$. However, from Figure 4 we can now see that, in the depicted range of MTTR increases, the relative increase in compensations is roughly on the same order as the MTTR increase. Consequently, the relative compensation increase is also on the same order as the relative reductions in MTTR-related expenses (Figure 2). In total, the inequality $q_\mathrm{R} > \psi \cdot q_\mathrm{C}$ is fulfilled even with a conservative assumption of $\psi \approx 0.1$, which means that reductions of combined expenses can be expected.

After this qualitative discussion, we consider the actual behavior of the reduction in combined expenses, q_T, as introduced in Section II. The combined expenses are composed of the MTTR-related expenses and the compensations, i.e., $E_\mathrm{T} = E_\mathrm{R} + E_\mathrm{C}$. In addition to the compensation policy, the MTTF, and the MTTR, the reduction in combined expenses also depends on the initial ratio of compensations to MTTR-related expenses, $\psi = E_{\mathrm{C},0}/E_{\mathrm{R},0}$. We use two exemplary values here: $\psi = 0.002$ and 0.1. In both cases, we assume the initial amount of compensation before the MTTR increase, $E_{\mathrm{C},0}$, to be $0.001 \cdot V$ (which is equivalent to $\mathbb{E}[C_0] = 0.001 \cdot \mathit{MRC}$). Consequently, for the conservative assumption of $\psi = 0.1$ we have $E_{\mathrm{R},0} = 0.01 \cdot V$ and for the more realistic assumption $\psi = 0.002$ we have $E_{\mathrm{R},0} = 0.5 \cdot V$.



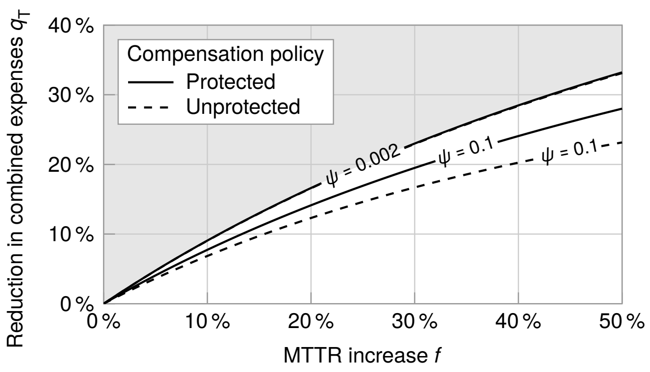
(a) Protected service.



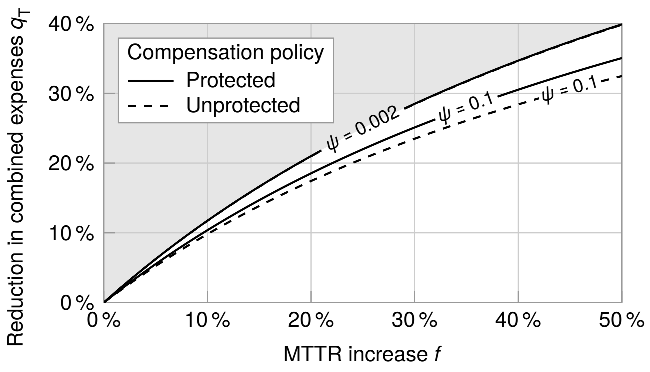
(b) Unprotected service.

Fig. 4. Relative increase in expected compensations for different initial MTTR values.

Figure 5a depicts the reduction in combined expenses for $\mathit{MTTR}^* = 0\,\mathrm{h}$ and $\mathit{MTTR}_0 = 5\,\mathrm{h}$, i.e., $\Phi = 0$. As can be seen, significant reductions are possible if the MTTR is increased. The curves show sublinear behavior, which means that initial MTTR increases are most effective. Nevertheless, also MTTR increases of more than $50\,\%$ lead to further reductions in combined expenses. With ψ as small as 0.002, the impact of the increase in compensations, q_C, on the combined expense decrease is almost negligible and, hence, q_T is dominated by the behavior of the decrease in MTTR-related expenses, q_R. Therefore, the behavior of the curves for the protected and unprotected compensation policy is nearly identical. On the other hand, for $\psi = 0.1$, the potential for expense reductions is lower and the compensation policies have a stronger impact. Since the compensation increase for an unprotected service is higher than for a protected one (Figure 4), the reduction in combined expenses is lower. The shaded area in Figure 5a



(a) $MTTR^* = 0\,\mathrm{h}$ and $MTTR_0 = 5\,\mathrm{h}$.



(b) $MTTR^* = 3\,\mathrm{h}$ and $MTTR_0 = 12\,\mathrm{h}$.

Fig. 5. Relative reduction of combined expenses. The shaded area represents $q_\mathrm{T} > q_\mathrm{R}$ and is not attainable. For $\psi = 0.002$, the curves of the protected and unprotected compensation policy are nearly identical.

represents $q_\mathrm{T} > q_\mathrm{R}$. Since q_R is the upper bound of q_T, the reduction in combined expenses must be below this area. As can be seen, the reductions for $\psi = 0.002$ are close to this bound.

Figure 5b shows the reduction in combined expenses for a scenario with $MTTR^* = 3\,\mathrm{h}$ and $MTTR_0 = 12\,\mathrm{h}$, i.e., $\Phi = 0.25$. The reduction in combined expenses is higher than in Figure 5a due to two reasons that add up. First and foremost, the ratio Φ is higher which leads to higher reductions in MTTR-related expenses (Figure 2). Second, $MTTR_0$ is higher which leads to smaller increases in compensations (Figure 4).

In summary, with a reasonable assumption for the ratio of initial compensations to MTTR-related expenses (e.g. $\psi \approx 0.002$), Φ (the ratio of minimal to initial MTTR) is the parameter with the most significant impact on the final expense reductions, because the behavior of q_T is dominated by that of q_R. For higher values of ψ, other parameters like the compensation policy, the initial MTTR or the MTTF gain in importance. In any case, MTTR increases as low as $10\,\%$ lead to reductions in combined expenses on the same order. Since the combined expenses represent a large share of the total expenses, *E*, and also the revenue, *V*, significant overall savings are possible for network operators.

VI. CONCLUSIONS

In this paper we have presented a potential approach for network operators to reduce their expenses and in that way to improve their profit margin. More specifically, we have suggested that network operators reduce their MTTR-related expenses and accept the associated increases in SLA compensation payments. To demonstrate the benefit for network operators, we have studied the trade-off between reductions of MTTR-related expenses and the increases in compensations using a detailed model. The model involves corporate key figures, a cost function that associates MTTR-related expenses with the MTTR, and a probabilistic estimation of the expected compensations. We showed that significant reductions for the sum of MTTR-related expenses and compensations are possible if the MTTR is increased. It should be noted that we used a ceteris paribus approach in our analysis, assuming e.g., no additional customer churn. If the MTTR is increased too much, this assumption is likely to be broken and additional effects would have to be considered. The extent of the overall reductions depends on various parameters of the network and the corporate key figures, but even with conservative assumptions, network operators can expect significant savings.

REFERENCES

[1] W. D. Grover and A. Sack, “High Availability Survivable Networks: When is Reducing MTTR Better than Adding Protection Capacity?” in *2007 6th International Workshop on Design and Reliable Communication Networks*. IEEE, 2007, pp. 1–7.

[2] C. Meusburger and D. A. Schupke, “Method to Estimate the Break-Even Point Between SLA Penalty Expenses and Protection Costs,” in *Optical Fiber Communication Conference*. Optical Society of America, 2008.

[3] U. Franke, “Optimal IT Service Availability: Shorter Outages, or Fewer?” *IEEE Transactions on Network and Service Management*, vol. 9, no. 1, pp. 22–33, 2011.

[4] Lumen Technologies, “2020 Annual Report and 2021 Proxy Statement,” https://s24.q4cdn.com/287068338/files/doc_financials/2021/ar/Lumen-2020-Annual-Report-and-2021-Proxy-Statement.pdf, 2021, Online, accessed 3 November 2021.

[5] Verizon Communications Inc., “2020 Annual Report on Form 10-K,” https://www.verizon.com/about/sites/default/files/2020-Annual-Report-on-Form-10-K.PDF, 2021, Online, accessed 3 November 2021.

[6] E. Hernandez-Valencia, S. Izzo, and B. Polonsky, “How Will NFV/SDN Transform Service Provider OpEx?” *IEEE Network*, vol. 29, no. 3, pp. 60–67, 2015.

[7] Lumen Technologies, “Lumen Service Level Agreement,” https://assets.lumen.com/is/content/Lumen/lumen-service-level-agreementpdf, 2021, Online, accessed 3 November 2021.

[8] CenturyLink, Inc., “CenturyLink Domestic Optical Wavelength Service; Service Level Agreement,” https://www.centurylink.com/legal/docs/QCC-Optical-Wavelength-Service-SLA.pdf, Online, accessed 3 November 2021.

[9] Verizon Communications Inc., “Wavelength Services Solution +; Part IV: Service Level Agreement,” https://enterprise.verizon.com/service_guide/reg/cp_wss_plus_sla.pdf, 2020, Online, accessed 3 November 2021.

[10] L. Takács, “On certain sojourn time problems in the theory of stochastic processes,” *Acta Mathematica Academiae Scientiarum Hungarica*, vol. 8, pp. 169–191, 1957.

[11] T. Enderle, “On the Impact of Billing Cycles on Compensations in Network SLAs,” in *2021 International Conference on Optical Network Design and Modeling (ONDM)*. IEEE, 2021, pp. 1–6.

[12] P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright *et al.*, “SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python,” *Nature Methods*, vol. 17, pp. 261–272, 2020.

[13] S. Verbrugge, D. Colle, P. Demeester, R. Huelsermann, and M. Jaeger, “General Availability Model for Multilayer Transport Networks,” in *Proceedings of the 5th International Workshop on Design of Reliable Communication Networks (DRCN 2005)*. IEEE, 2005, pp. 85–92.