How to Use Mobile Communication in Critical Infrastructures: A Dependability Analysis

Jonas Wäfler(✉) and Poul E. Heegaard

Norwegian University of Science and Technology, 7491 Trondheim, Norway
{Jonas.Waefler,Poul.Heegaard}@item.ntnu.no

Abstract. Critical infrastructures, like the future power grid, rely strongly on a reliable communication infrastructure. Mobile communication seems an attractive candidate, as the entry costs are low and, provided the coverage, the new devices have immediate communication access upon installation. However, considering the long time-frame of this investment, it is important to think about the constraints in mobile networks and also potential challenges waiting in the future. In this study, which is based on the situation in Norway, we discuss four important future challenges: policy change, contract change, change of *Quality of Service* and network failure. We show that a clever use of mobile communication like multihoming or using a mobile virtual network operator may meet the challenges. In the second part, we quantify the availability of the different mobile communication usages with the help of analytical models and show that already a small increase of additional battery capacity in the mobile network improves the availability significantly.

Keywords: Mobile communication · Critical infrastructure · Battery backup · Smart grid · Availability · Interdependencies · Markov model

1 Introduction

Like other critical infrastructures, the future power grid is going to rely strongly on a reliable communication infrastructure. Intelligent electronic devices (IED) are going to be deployed throughout the power grid and are in need of a flexible communication platform [1]. The requirements concerning latency, availability and security [2,3] are very diverse and might be covered by either a flexible middleware framework for data communication like GridStat [4] or a mixture of different technologies. Among the considered technologies, mobile communication is regarded as a pragmatic choice for services like smart metering and monitoring in remote locations. It is a tempting candidate, because the entry costs are relatively low and, provided adequate coverage, the device has immediate communication access upon installation. However, there are many pitfalls to avoid, not least because of the long term nature of the investment.

The mobile networks conduct an access control based on the mobile device's subscription. A device is usually only allowed to use the network of the operator,

F. Koornneef and C. van Gulijk (Eds.): SAFECOMP 2015 Workshops, LNCS 9338, pp. 335–344, 2015.
DOI: 10.1007/978-3-319-24249-1_29

which issued the subscription. National roaming, i.e. the communication over networks of other operators, is technically possible but commonly not permitted. There are exceptions for special numbers like police and fire department and for special groups of customers, e.g. in Norway the regulator stipulated national roaming for a limited set of prioritized customers from rescue organizations [5]. If a utility wants to use a different operator because the reception has deteriorated or it changed the contract, it has to manually exchange the SIM card in the device, which may be very costly as the potential number of devices for smart metering and monitoring is very large.

An important property for the suitability of a communication infrastructure is its dependability. Only few public studies exist [6–8] as the access to data is usually restricted. The first two studies focus on operator internal incidents, the third one [8], however, takes a different approach: it is based on measurements done by mobile devices distributed over 300 different places in whole Norway. The logged connectivity to the different UMTS networks show the distribution of time between failures, down time and unavailability. This study measures the Quality of Service exactly how a user would perceive it.

In this paper we suggest several alternatives on how a power utility may use mobile communication; we single out the four main future challenges and analyze how the alternatives react to those. After this qualitative analysis we analyze the availability of the alternatives quantitatively based on measurement data from the study from [8]. And finally, we analyze the availability improvement when equipping the base stations in the mobile network with more battery capacity.

2 System Description

We consider the case, in which a company wants to roll-out a large number of mobile devices. These devices could be smart meters or monitoring devices inside the power grid. The study focuses on the implication of using mobile communication for these smart devices, this is done by concentrating on the communication between a single smart device and the company. The mobile communication is provided by two mobile network operators (MNO): *MNO A* and *MNO B*. It is assumed, that there is no national roaming agreement between *MNO A* and *MNO B*, i.e. subscribers of one network have no access to the other network. As in real networks, the two infrastructures are not completely independent and thus their failures manifest some dependencies. The reason is twofold. First, shared infrastructure or geographical collocation of infrastructure in certain parts of the network, e.g. A leases a communication line from B in rural and sparsely populated areas or A and B have their cables in the same ditch. Second, dependence on the same service like for example power supply. In both cases one failure can cause a failure in the two MNOs.

The MNOs are considered as black boxes, no internal state is known, the mobile device only knows whether a connection to an MNO is possible and, on a higher network level, if it has a connection to the power utility. It is assumed, that only the MNOs can fail, as they are the main focus of the study.

In order to connect to the mobile network any device needs a SIM card. On each SIM card there is a number (IMSI) which uniquely identifies each device. Part of this number is the mobile network code (MNC), which identifies the mobile company that issued the SIM card. Access control is based on the MNC, an MNO allows only connections from devices with its own MNC or with an MNC belonging to an MNO with a roaming agreement. In Norway, these roaming agreements are scarce and limited to foreign MNOs or mobile companies owning no or only a very limited network on their own.

2.1 Challenges

Any mobile solution faces challenges over its lifetime. In the following we list the challenges, which are in our opinion the most important once.

Challenge 1: Policy Change. Mobile communication depends on policies from the national regulator and also on policies from the MNO. The national regulator may for example forbid international roaming fees or impose national roaming; the MNO may change national and international roaming agreements.

Challenge 2: Contract Change. The contract between the subscriber and the MNO is subject to changes over time. Examples are an increase of the subscription fee above an acceptable price level, required services that are discontinued, bankruptcy of the MNO or its acquisition.

Challenge 3: Change of QoS. The *Quality of Service* (QoS) at a device may change over time. Examples are a reduced signal strength or increased blocking probability because of structural changes between the mobile device and the base station (e.g. new walls, new buildings) or changes in the usage pattern of the base station (e.g. increased number of subscribers).

Challenge 4: Network Failure. A network failure in this context is defined as service outage, i.e. communication from sender to receiver over this specific network is not possible. The mobile device always tries to connect to a base station of its prioritized MNO. If no base station of its prioritized MNO is available, it may try to connect to a base station of another MNO, but a connection is only established if a roaming agreement with that MNO exists.

The time granularity is very different and decreases from the first to the last challenge, i.e. the reaction time for the operator is getting shorter. Policy and contract changes have to be announced with a certain lead time and the operator can look for a solution well in advance. A change of QoS, however, may happen without notice and network failures usually come without warning and the system has to immediately react to mitigate the failure.

3 Usage Alternatives

The ordinary way is to buy regular SIM cards from an existing MNO, denoted in the following as *ordinary subscription*. This comes with a carrier lock-in:

a change of MNO can only be achieved by replacing the SIM card in each and every device. This is costly, as the number of devices is likely to be high and some of the devices may be located in remote areas or in places difficult to reach. Also a network failure has a strong impact, as a national roaming is usually not allowed, i.e. only the network of your own MNO can be used.

MVNO. The utility takes the role of a *mobile virtual network operator* (MVNO), buying a certain amount of services from an MNO. Utilities may collaborate nationally to reduce the operational costs.

The MNO can be changed by changing roaming agreements. There are already many MVNOs, so this is a proven solution and it can be implemented quickly by out-sourcing almost everything if desired. A precondition for this solution is that existing MNOs allow roaming by MVNOs. A policy change by the national regulator or the MNOs may therefore have an impact on this solution. An MVNO has usually only an agreement with one MNO and it may happen that no MNO can provide a satisfactory QoS for all the devices. In this case, changing the MNO does not help. This threat is higher for geographically wide spread utilities. In case of a network failure, this solution has the same weakness as the *Ordinary Subscription*, because the network cannot be changed on short notification but needs longer negotiations.

The MVNO may issue several series of SIM cards with different MNCs. It can then make individual roaming agreements for each MNC. This way some of the discussed problems can be mitigated.

Multihoming. Certain devices allow the use of multiple SIM cards. Using a SIM card from each MNO implements a national roaming without dependencies on policy changes by the regulator or the MNOs. An application on the device probes the different networks and chooses the one with the most favorable QoS. There is a carrier lock-in, however, by using several SIM cards the risk is minimized. Using a SIM card from an MVNO especially for utilities may increase the flexibility of this solution even more. A new MNO can only be used by inserting their SIM card. The cost per device is higher, as it needs multiple SIM card slots and multiple subscriptions per device.

International Subscription. Interestingly, users with a foreign subscription can have an advantage over those with a national subscription when the foreign MNO has roaming agreements with several national MNOs. In this case, the foreign subscription implements a national roaming.

The advantages are that it is very easy to implement and several mobile networks can be used, depending on the roaming agreements. The switchover to another network may be fast, depending on the network failure. International roaming depends strongly on the policies of the regulator and the MNOs that are in place. If the roaming costs are abolished for good, the MNOs may restrict roaming agreements or make international coalitions with roaming agreements.

But all depends strongly on what is de fined as legal by the European and the national regulator. Additionally, this solution leads again to a carrier lock-in.

4 Unavailability

The availability of the alternatives can be grouped in three classes.
$\boldsymbol{A}_{\mathbf{single}}$: only one single network is used, if it fails the connection fails as well;
$\boldsymbol{A}_{\mathbf{standby}}$: there is a standby network, which is used in the case of a failure in the primary one, the switchover time varies between the solutions;
$\boldsymbol{A}_{\mathbf{DMR}}$: (DMR: dual modular redundancy): two networks are used at the same time and a failure in one does not interrupt the connection.

The *ordinary subscription* and *MVNO* (with one MNC) are in the class A_{single} because they can only use the network of a single MNC, namely the one having issued the SIM card or the one having a roaming agreement, respectively. The solution *MVNO* (with multiple MNCs) is either in the class A_{single} or A_{standby}, depending on whether the MNC is fix or whether it can be changed dynamically in case of a network failure. *Multihoming* is in the class A_{DMR} if the SIM cards are used in parallel and in class A_{standby} if one is in a standby state. The *international subscription* is in the class A_{standby} because the device can only be connected to one network at a time and needs to reconnect in the case of a network failure.

We compute the unavailability U of the classes, given by $U = 1 - A$, where A is the availability defined as "*readiness for correct service*" [9].

4.1 Quantification of $\boldsymbol{A}_{\mathbf{single}}$ and $\boldsymbol{A}_{\mathbf{DMR}}$

The mentioned study [8], contains data for our classes A_{single} and A_{DMR}. Additionally, it also contains the distributions for *time between failure* and *down time* when using a single network. Assuming the distributions to be negative exponential, the failure and restoration rates are computed with the approximated *mean time between failure (MTBF)* and *mean down time (MDT)* by $\lambda = 1/(\text{MTBF-MDT})$ and $\mu = 1/\text{MDT}$. The parameters are given in Table 1. The two networks have very different properties: *MNO A* has more failures than *MNO B*, but due to its short restoration time it has a lower overall unavailability.

Table 1. Used parameters from study [8].

	Unavailability U	Failure rate $\lambda_{i,total}$ [s^{-1}]	Restoration rate μ_i [s^{-1}]
A_{single}	3.3×10^{-4}	1.11×10^{-5}	3.33×10^{-2}
A_{single}	5.0×10^{-3}	2.01×10^{-6}	4×10^{-4}
A_{DMR}	2.0×10^{-5}	–	–

4.2 Quantification of $\boldsymbol{A}_{\mathbf{standby}}$

There are no numbers for A_{standby}, however, we show how it can be computed with a Markov model and the given parameters. But first, we note, that the measurements in Table 1 indicate, that *MNO A* and *MNO B* are *not* independent, they are subject to common cause failures. In order to compute this common cause failure rate the Markov model in Fig. 1 is used. The round states are system up states and the square states system down states. The state of the whole system is defined by the states of the two MNOs ($i_A : i_B$) with $i_A, i_B \in \{\text{ok,d,cf}\}$. The states for each MNO are working (ok), down (d) or down because of a common cause failure (cf). Common cause failures from states other than *(ok:ok)* are omitted for the sake of readability; the introduced error is negligible, as the *ok:ok* state has by far the highest state probability. The λ_is are computed by $\lambda_i = \lambda_{i,total} - \lambda_{cf}$ in order to keep the total failure rates $\lambda_{i,total}$ constant when varying λ_{cf}. Setting $\lambda_{cf} = 0$, i.e. making the networks independent, we get an unavailability of 1.67×10^{-6}, i.e. around 12 times smaller than the measured unavailability in Table 1, showing that the networks are in fact dependent as mentioned above.

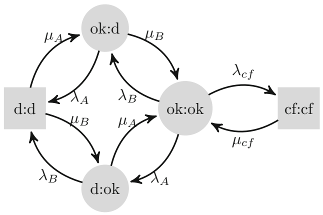


Fig. 1. Model for class A_{DMR}

Details about shared infrastructures and services in *MNO A* and *MNO B* are not known. However, leased line and power incidents are possibly large contributors to failures [6], therefore, we assume a restoration time of $\mu_{cf} = 2500s$, which is in the order of a longer mobile restoration time and a power outage restoration [10]. Solving the model with the unavailability and rates given in Table 1 yields a common cause failure rate λ_{cf} as listed in Table 2. The failure rate λ_{cf} makes around 5 % of the total failure rate of *MNO A* $\lambda_{\text{A,total}}$ and around 30 % of *MNO B* $\lambda_{\text{B,total}}$.

Table 2. Common cause rates after parameter fitting.

λ_{cf} [s^{-1}]	μ_{cf} [s^{-1}]
6.34×10^{-7}	4×10^{-4}

Finally, the unavailability for A_{standby} is computed by extending the state definitions to ($j_A : j_B$) with $j_A, j_B \in \{\text{ok, OK}, d, D, \text{cf, CF}\}$, which yields the model depicted in Fig. 2. Uppercase letters indicate that the mobile device is currently using that network. E.g, state ($ok : D$) means network B is used, but *down* and network A is *ok*. It is a down state (square), only after switching the network, leading to state ($Ok : d$) is the system up and running again.

In a business oriented setting it can be advantageous to prefer one MNO over the other because of special price models based for example on data volume. The other MNO is only used if the preferred one is down. For that, the model in Fig. 2 is adjusted to always switch over to the preferred network if it is working. i.e. if *MNO A* is preferred, adding a new transition from *(ok:OK)* to *(OK:ok)* and marking the former state as down state because of the unavailability during the switchover.

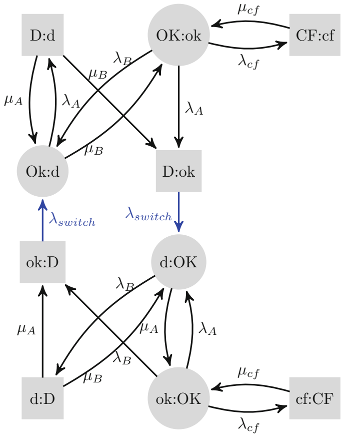


Fig. 2. Model for class A_{standby}

4.3 Discussion

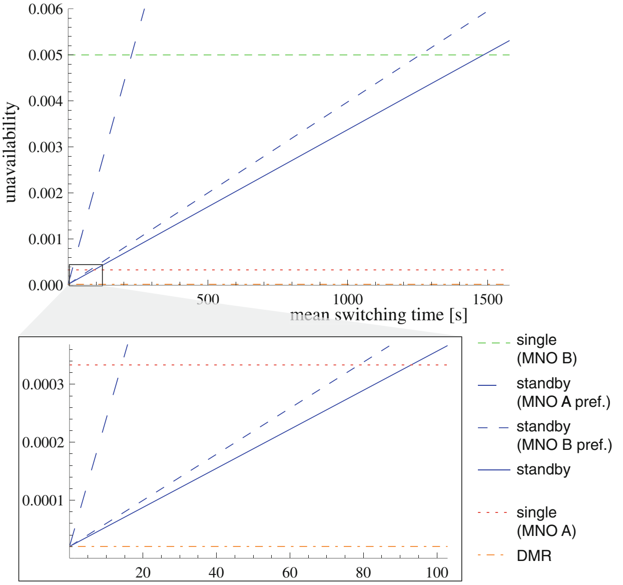


Fig. 3. Unavailability of the classes against switching time. Zoom-in for low values of switching time.

The results of a steady-state analysis are given in Fig. 3. They show clearly the large difference in unavailability of the different solutions. Class A_{single} has two results depending on which MNO is chosen. The difference between the two MNOs is big because of the large difference in restoration time.

In the class A_{standby}, the unavailability is linearly increasing with the mean switching time. The unavailability is lower than the unavailability of A_{single} if the mean switching time is lower than 95 s or 1485 s for *MNO A* and *MNO B*, respectively. The first number is surprisingly small, it is explained by the very short average restoration time in *MNO A* of $1/\mu_A = 30\,s$. The switching time itself depends strongly on the used alternative and implementation. Two alternatives belonging to the class A_{standby} may, therefore, not necessarily have the same unavailability.

Preferring one MNO leads to a higher unavailability. *MNO B* is here the better choice of the two, as this solution benefits from the longer uptime of *MNO B* and the shorter restoration time of *MNO A*. Preferring one MNO creates additional interruptions, i.e. a lower mean time between failure (MTBF) and should be avoided. However, as stated above there might be other considerations that need to be taken into account. We consider the system as down during the switchover, if it is performed without downtime, then preferring *MNO B* has a lower unavailability than the standard standby class.

5 Improving Availability with Batteries

Today, batteries are available in some base stations. Depending on the MNO the number of equipped base stations as well as capacity varies strongly. In Norway there are discussions between the national regulator and MNOs about stipulating a required battery installation in base stations in mobile networks [11]. So far, installed batteries in the power grid were already included implicitly, because we used measurements of actual networks. In the following we study the effect of installing additional battery capacity.

Batteries allow the communication system to keep on working in case of a power failure, if it is bridgeable by battery. We assume that this is the case for p % of all failures, valid for both individual failures and common cause failures. The battery capacity is assumed to be negative exponentially distributed with mean $1/\lambda_{\text{bat}}$. This assumption is justified by the variation of capacity due to different battery types, battery ages, working conditions and charging states.

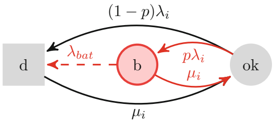


Fig. 4. Model for class A_{single} with limited battery capacity.

The extended models for the classes A_{single} and A_{DMR} are depicted in Figs. 4 and 5. The state definition is extended by the network state b, indicating that the network suffered a power failure and parts of it is running on battery. The dashed arrows indicate a transition caused by battery depletion. The model for A_{standby} is not depicted but is constructed as before by duplicating the model for A_{DMR}, adding an indication for which MNO is active and adding two new transitions with rate λ_{switch} between *ok:D* to *OK:d* and *D:ok* to *d:OK*.

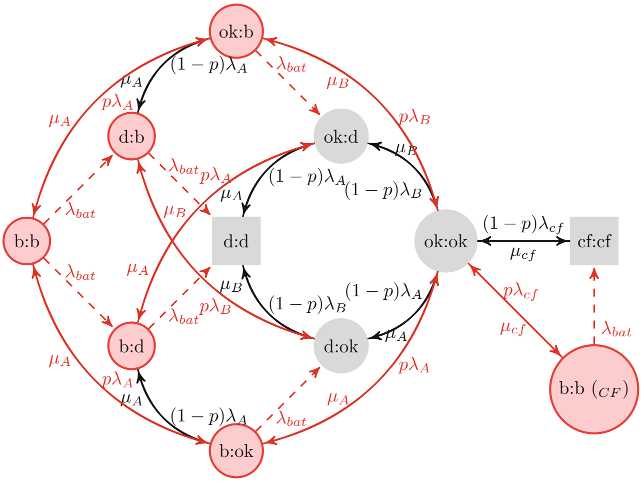
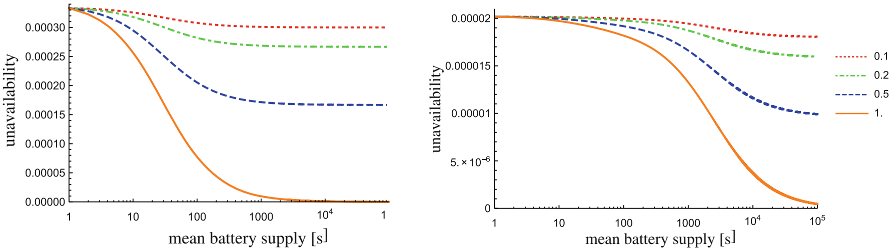


Fig. 5. Model for the class A_{DMR} with limited battery capacity.

5.1 Discussion

Figure 6(a) shows the results for the class A_{single} when using *MNO A*. The unavailability is most sensitive to a mean battery capacity in the order of the mean down time, i.e. $1/\mu_A = 30$ s. For the *MNO B* the plot would look similar, but shifted towards its mean down time of $1/\mu_B = 2500$ s.

Figure 6(b) shows the results for the class A_{DMR}. The two parameters λ_{cf} and μ_{cf} are set to the values used previously, noted in Table 2, which equals to a mean *common cause restoration time* of 2500 s. As expected are the absolute values lower than in the class A_{single}; the plot is in fact almost the same as for *MNO B*, except the y values are much lower. The reason being, that of the two down states in the model, the state *cf:cf* is responsible for the highest fraction



(a) class A_{single} (b) class A_{DMR}

Fig. 6. Unavailability against battery capacity for different values of p.

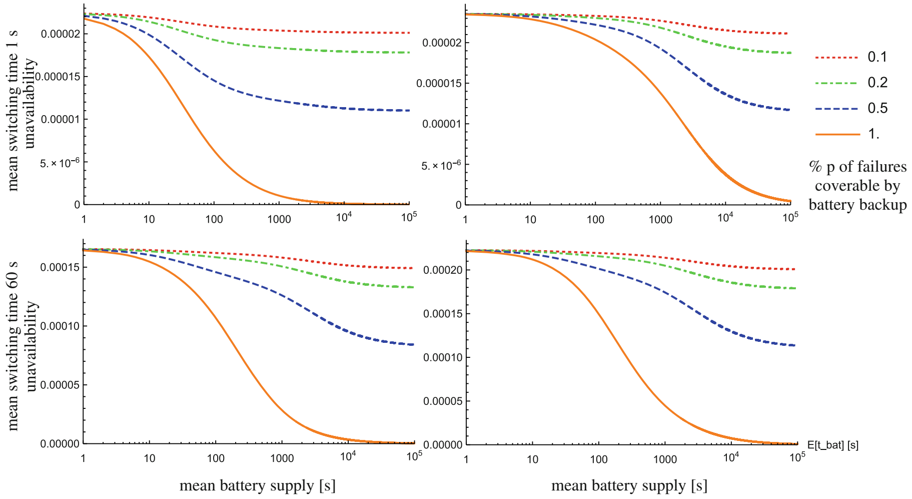


Fig. 7. Unavailability vs battery capacity for class A_{standby} with different values of p.

of the down time. The mean sojourn time for this state is given by $1/\lambda_{cf}$ and is equal to the restoration time in *MNO B*.

Figure 7 shows the results for the class A_{standby}. The simulation is done for two scenarios with different pairs for λ_{cf} and μ_{cf}. In scenario 1, $1/\mu_{cf}$ is chosen to be very short, i.e. 30 s, which corresponds to the restoration rate of *MNO A*. As before, λ_{cf} is given indirectly by the model in Fig. 1 by solving the steady state equations for it. In scenario 2, the two parameters λ_{cf} and μ_{cf} are set to the values used previously, i.e. $1/\mu_{cf}$ of 2500 s. Additionally, it is done for two different switching times. For a switching time of 1 s the difference between the two scenarios is big, i.e. the downtime caused by the common cause failure is dominant. When increasing the switching time to 60 s, however, the downtime caused by the switching itself becomes dominant and the difference between the two scenarios is minimal.

The numbers show that the availability gain can already be large for a small battery capacity bridging a time of 1–3 min. However, it depends strongly on the restoration times and switching times between the networks.

6 Conclusion

We list different alternatives of how to use mobile communication in this paper. By combining them, more are possible, but they are not fundamentally different to the presented ones. As the machine-to-machine communication (M2M) is likely to increase in the future, new technologies and especially new regulations may change the way mobile communication is used. For example, a decoupling of the SIM card and the operator by issuing carrier-free SIM cards would allow the switching between different networks and subscription contracts with only a short switching delay. This would inexpensively implement a virtual multihoming belonging to the availability class A_{standby} as discussed above.

This study is based on the regulation status and availability statistics in Norway. Details might be different in other countries. If and how mobile communication should be used depends on what service is run over it and its requirements concerning availability, performance and costs. In this paper we only focused on future challenges, usage alternatives and the availability; performance and costs are important factors but were outside the scope.

References

1. International Energy Agency (IEA), Technology roadmap: Smart grids (2011)
2. Bakken, D.E., Bose, A., Hauser, C.H., Whitehead, D.E., Zweigle, G.C.: Smart generation and transmission with coherent, real-time data. Proc. IEEE **99**(6), 928–951 (2011)
3. Electric Power Research Institute (EPRI), The integrated energy and communication systems architechture, vol. IV, Technical analysis, Technical Report (2004)
4. Gjermundrod, H., Bakken, D.E., Hauser, C.H., Bose, A.: GridStat: a flexible QoS-managed data dissemination framework for the power grid. IEEE Trans. Power Deliv. **24**(1), 136–143 (2009)
5. Forskrift om prioritet i mobilnett, [Regulation about Priorities in Mobile Networks], FOR-2013-10-21-1241, NKOM, Norwegian Communication Authority (2013)
6. Følstad, E.L., Helvik, B.E.: Failures and changes in cellular access networks: a study of field data. In: Proceedings of DRCN, pp. 132–139 (2011)
7. Matz, S.M., Votta, L.G., Malkawi, M.: Analysis of failure and recovery rates in a wireless telecommunications system.: In: Proceedings of Dependable Systems and Networks (DSN), pp. 687–693 (2002)
8. Kvalbein, A.: Robusthet i norske mobilnett, [Robustness in Norwegian mobile networks], simula research laboratory, Technical Report (2013)
9. Avizienis, A., Laprie, J.C., Randell, B., Landwehr, C.: Basic concepts and taxonomy of dependable and secure computing. IEEE Trans. Dependable Secure Comput. **1**(1), 11–33 (2004)
10. Avbrotsstatistikk 2013, [Outage statistics 2013], NVE, Norwegian Water Resources and Energy Directorate (2014)
11. Sikkerhet og beredskap mot ekstremvær i telesektoren, [Security and preparedness for extreme weather situations in the telecommunication sector], Working Group Energi Norge and Telenor, Technical Report (2013)