Design and initial validation of the Raster method for telecom service availability risk assessment

Eelco Vriezekolk
Radiocommunications Agency Netherlands
& University of Twente
eelco.vriezekolk@agentschaptelecom.nl

Roel Wieringa
University of Twente
roelw@ewi.utwente.nl

Sandro Etalle
Eindhoven University of Technology & University of Twente
s.etalle@tue.nl

ABSTRACT

Crisis organisations depend on telecommunication services; unavailability of these services reduces the effectiveness of crisis response. Crisis organisations should therefore be aware of availability risks, and need a suitable risk assessment method. Such a method needs to be aware of the exceptional circumstances in which crisis organisations operate, and of the commercial structure of modern telecom services. We found that existing risk assessment methods are unsuitable for this problem domain. Hence, crisis organisations do not perform any risk assessment, trust their supplier, or rely on service level agreements, which are not meaningful during crisis situations. We have therefore developed a new risk assessment method, which we call RASTER. We have tested RASTER using a case study at the crisis organisation of a government agency, and improved the method based on the analysis of case results. Our initial validation suggests that the method can yield practical results.

Keywords

Risk assessment, telecommunication services, availability.

INTRODUCTION

Crisis organisations (such as police, fire services, and emergency medical care) depend on telecom services. Without telecommunication services, they cannot coordinate their operations, receive situation updates, or inform stakeholders. Unavailability may be due to cable breaks, interference, congestion, power failures, operator mistakes, and many other technical and non-technical causes. Failure of telecommunication services during the response phase will likely increase damage and could cost lives. Many crisis organisations have adopted net-centric operations, thereby increasing their effectiveness but also increasing their vulnerability to the unavailability of telecom services.

Crisis organisations should be aware of their telecom service availability risks. This means that they should perform a risk assessment (RA). The goal of a RA is to understand and evaluate all risks, and to provide the basis on which risk treatment decisions can be made. Unfortunately, RA in this domain is challenging. This is mainly due to two issues: the special circumstances in which crisis organisations operate, and the way modern telecom services are organised.

On the first issue: crisis organisations operate during disasters and crises, outside the "normal" conditions for which telecom services are primarily designed, and on which service level agreements are based. They also operate under close public scrutiny; as members of society we all benefit from professional crisis response and we are therefore all stakeholders. Where most telecom users work toward economic goals, crisis organisations aim to satisfy public goals.

On the second issue: modern telecom services are delivered using a composition of independent infrastructures, owned and maintained by independent, often competing companies. Even a simple phone call travels through several private and public networks. This composition requires that constituent infrastructures are decoupled; they do not need to know the details of the services delivered by their neighbouring networks. By implication, telecom providers typically do not know in which end-to-end services they participate. Their own internal risk

priorities must therefore be based on general scenarios, and may not be suitable to the needs of crisis scenarios. It also implies that end-users will find it difficult or impossible to gain detailed insight in the workings of the infrastructures on which they are dependent.

In previous work we briefly presented requirements on RA methods in this domain, and sketched our initial ideas for a new RA method (Vriezekolk, Wieringa and Etalle, 2011a). These requirements were elicited from the risk assessment practice in the organization of the first author, Radiocommunications Agency Netherlands; responsibilities of the Agency include the availability of telecom infrastructures during incidents and disasters in the Netherlands. One of these requirements concerns the inclusion of a diverse range of risk factors, which we investigated in (Vriezekolk, Wieringa and Etalle, 2011b). The present work extends these previous papers. In particular, we have designed a RA method to satisfy the requirements, which we call RASTER. To show that RASTER does work in practice and to validate some of its requirements, we have done a case study. We learned a lot from this case study, and improved RASTER based on these findings. After the case study, we asked the participants to complete a questionnaire and interviewed them, to help us validate some of our assumptions and requirements. Our experiences with RASTER will be of interest to other RA practitioners as well as to crisis responders who are responsible for telecom issues.

The structure of this paper is as follows. First, we give an overview of existing approaches to RA. We then motivate our requirements, and compare them to existing approaches. Our research method is described next; we introduce the RA method through an incremental improvement process, in which we present a version of the method, test it, and improve it based on the results of the test. We iterated four times through this improvement process. Results of the case study and questionnaire are presented and discussed, and we conclude the paper with implications for further research.

RELATED WORK

In an earlier paper we explored nine requirements to methods for risk assessment in our domain, numbered R1 to R9 (Vriezekolk et al., 2011a). These requirements are based on nine years of work experience of the first author at Radiocommunications Agency Netherlands. This is a Dutch government agency responsible for operational activities and enforcement of regulations pertaining to telecommunications.

In its operations, the Agency is confronted with society's increased dependency on telecommunications services. This is evidenced, for example, by the increasing scarcity of available space in the parts of the radio spectrum that are suitable for telecommunications (Radiocommunications Agency Netherlands, 2009 and 2010). As a result, two issues become more important. First, there is great interest in incidents that cause outage, and telecom risks in general. Second, and because of this interest, explanations on policy decisions and justification of the Agency's actions are addressed no longer exclusively to experts, but increasingly to non-experts, such as politicians and concerned citizens (Vriezekolk et al., 2011b). Furthermore, private and public sector organisations alike have a pressing need to reduce costs; there are few resources for in-depth studies, unless they yield results that can be applied immediately. For these various reasons, the Agency can no longer limit itself to isolated elements of telecom systems, such as the radio path only, but has to consider the delivery of telecom services from end user to end user.

The nine requirements listed in the referenced paper derive from the trends described in the previous paragraph. Specifically, this relation is as follows. Requirements R1 to R3 state that, respectively, the method must gracefully deal with missing information, with uncertainty, and with the reactions of stakeholders to risk treatment decisions (which we call 'social risk factors'). These requirements derive from the involvement of non-experts, and characteristics of the telecoms market and its regulation. Requirements R4 to R6 state that a team with analysts from different disciplines must be able to execute the method (R4) in a (to them) reasonable amount of time (R5), taking account of the incidents and disasters that occur when crisis organisations are active (R6). These requirements derive from the mix of disciplines involved in crisis management, and the need for budgetary restraint. Finally, requirements R7 to R9 state that the method must provide the decision makers with assurance that all threats have been identified (R7), clear risk treatment options (R9) and, more importantly, with justifications for them (R8). These requirements derive from the heightened interest in incidents and risks, especially by non-experts.

We see two sources from which telecom risk assessment methods originate. The first is safety assurance and accident investigation for general complex technical systems, such as nuclear installations, industrial plants or telecommunications. The second comprises methods tailored for information security for information systems.

Safety assurance and accident investigation do not have a very long history. Hale and Hovden have identified three ages of industrial safety (Hale and Hovden, 1998). Each age contributed new important insights and techniques to the previous state of the art. The first age, the technological age (ca. 1920s), was dominated by the

technical aspects of risk. Well known techniques such as Fault Tree Analysis (FTA), Failure Modes and Effects Analysis (FMEA) and Hazard and operability studies (Hazop) were developed in this age (IEC/ISO standards 60812, 61025, 61882). Increasing awareness of human factors in risk and safety led to a new wave of risk management techniques, forming the second age (ca. 1970s). Methods such as Technique for Human Error Rate Prediction (THERP) and Human Error Assessment and Reduction Technique (HEART) belong to this period (Swain and Guttmann, 1983; Williams, 1988). The third age of safety (ca. 1980s) expanded the focus to organisational safety culture, with Management Oversight and Risk Tree (MORT) and TRIPOD as representative examples (Johnson, 1973; Reason, Shotton, Wagenaar, Hudson and Groeneweg, 1989). Each stage led to improved and more complex risk management techniques. But despite enhanced risk management, Perrow showed that avoidable accidents still happened (Perrow, 1984). While Perrow has argued that in systems with complex interactions and tight coupling accidents are inescapable and hence "normal", LaPorte has identified organisations whose safety culture made them into "high-reliability organisations" (LaPorte and Consolini, 1991). More recently, accident models and risk management techniques have emerged that view safety as an emergent property of the interactions within complex systems. Notable examples are STAMP and FRAM (Leveson, Dulac, Marais and Carroll, 2009; Hollnagel and Goteman, 2004).

Interest in safety and security for information systems is, of course, of much more recent time than for industrial installations. Human factors have been a focus of concern from the beginning, but mainly in the form of vulnerability to malicious attacks. Not only availability, but also confidentiality and integrity of information and systems are prime objectives. Well-known risk assessment methods are the CCTA Risk Analysis and Management Methodology (CRAMM) and the CORAS framework (Barber and Davey, 1992; Fredriksen, Kristiansen, Gran, Stølen, Opperud and Dimitrakos, 2002).

Existing RA methods do not fully address the combination of incomplete information about a complex system, large uncertainties, and social risk factors. Since we lack detailed information on the telecom infrastructures (requirement R1), methods such as FTA cannot be applied. The analytical parts of the FMEA method can be of use but, again because of requirement R1, we cannot easily decompose the system (the end-to-end telecom service) into layers and elements as required by FMEA. Highly quantitative methods, such as Event Tree Analysis, cannot be used because accurate data is not available (requirement R2). Such methods would not be sufficient either, because in our domain qualitative risk factors are relevant as well (requirement R3). Methods such as STAMP and FRAM are complex (requirement R4) and time consuming (requirement R5). Requirement R6 implies that service level agreements (SLAs) are not a sufficient basis by themselves, as service level agreements typically exclude disasters and other exceptional circumstances.

Methods designed for information systems, such as CRAMM and CORAS, are not suitable either. These methods typically model the information dependencies between components, for example modelling the fact that a front-end system depends on information in a customer database. Such relationships are not relevant to us, since end-users do not store or retrieve information inside the telecom service. To us, telecom services have the appearance of an industrial installation, not of an information system.

Current practice also does not reveal any suitable RA methods. Most organisations are not aware of their availability risks at all. Some rely on the professionalism of their provider, or on an agreed SLA (which is often insufficient and cannot replace risk assessment, as was shown above).

We believe that this sufficiently justifies our development of an alternative RA method, one that is tailored to our problem domain. Our proposed method is not entirely new. It contains elements from FMEA, with extensions to handle incomplete and missing information, and social risk factors.

RESEARCH METHOD

We designed a first version our new RA method, which we expected would satisfy the requirements. We validated this initial version in a case study and redesigned it based on initial case experience. Validation was continued with the improved RA method. This process was repeated two more times, continuing the case study every time with an improved version of the method. This methodology of incrementally designing and validating a technique is described in more detail by Wieringa (2009). To evaluate the case study, we asked the participants to complete a questionnaire and conducted a final interview with the participants.

Our only research question at this stage is whether the method can be practically applied. Validations of its usefulness and completeness of the results will be done in later case studies, as these require a workable method first. We operationalize 'applicability' by three questions: 1) Are the analysts able to apply and complete the method based on the manuals provided to them? 2) Does the method yield explicable treatment recommendations? 3) Are the analysts satisfied with the results and the time needed to complete the method?

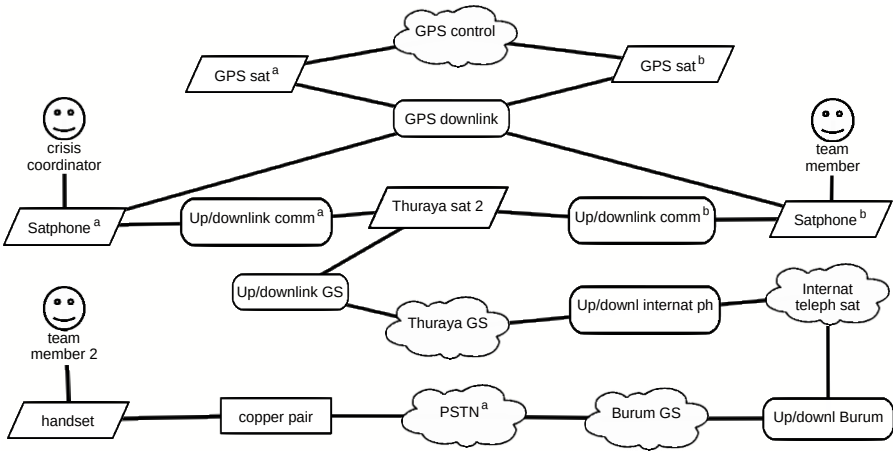

Figure 1: Example of a model, showing a satellite telephone system connecting to the international telephony system.

INITIAL DESIGN OF THE METHOD

Models of telecom services are central to the method. These models consist of a graph that represents the physical components that are necessary to deliver end-to-end communication. Each node of the graph is typed as an actor, wireless link, wired link, equipment item, or unknown link. In diagrams, the type of a node is indicated by its shape. Figure 1 shows an example that models a satellite telephony system. Two actors each use a satellite phone (the parallelogram marked "Satphone"). Wireless links (indicated by rounded rectangles) provide the link to a satellite ("Thuraya sat 2"). Further components model the dependency on the GPS signal, and connectivity to the terrestrial telephone system. Unknown links (cloud shapes) represent areas of the infrastructure for which no information is available, or for which it is not necessary to know its detailed makeup.

When analysts find that they lack the information to reliably assess threats for a node, this is a trigger to perform further research. In case of unknown links, it signifies that the internal structure of the unknown link needs to be exposed to the model. Since the method starts with a high-level model and adds detail only as and when required, we called it Risk Assessment by Stepwise Refinement (RASTER).

We now present the four stages of the RASTER method. The method closely follows the standard break-down of RA into preparation, risk identification, risk analysis, and evaluation (ISO, 2009). However, as a result of stepwise refinement, RASTER executes risk identification and risk analysis in parallel. The four stages of RASTER are depicted in Figure 2.

The *first stage* (initiation and preparation) determines the boundaries and assumptions of the RA. The information prepared during this stage will be used in further stages. This information includes the goals and operational procedures of the crisis organisation, likely crisis scenarios, and a complete list of all telecom services used by the organisation.

In the *second stage* (single failures analysis), threats to individual nodes are analysed. The method provides checklists of common threats, but analysts are encouraged to add specific threats to each node. Analysis includes a qualitative estimate of the likelihood of the threat leading to unavailability of the node, and the magnitude of that impact. These scores only take risk factors into account that we call physical, to contrast them with treatment factors and attitudes, mentioned later: likelihood, uncertainty, magnitude and geographical and temporal extent of the damage, and repairability (persistency).

Impact is assessed on a scale with seven classes. Three basic classes are used: low, medium, and high. As indicative description for a 'low' impact we use "noticeable degradation of the service". A 'medium' impact is described as "partial, temporary unavailability of the service". A 'high' impact is described as "complete, prolonged unavailability of the service". In addition, two classes are used to indicate a likelihood that is extremely high or low. The extreme classes differ from the basic low/medium/high classes; they indicate cases where regular quantitative reasoning ceases to be meaningful, for example when an impact is so high that it is deemed unacceptable regardless of its likelihood. In situations when insufficient data is available to make a reasonably accurate assessment, the class 'unknown' is used. Finally, the class 'ambiguous' is used when analysts differ in their estimates and cannot reach consensus on a common assessment. Ambiguity is thus treated separately from 'normal' uncertainty.

Frequency is assessed on a similar seven-class scale. As indicative values for 'low' a likelihood of once per 500 years is used; for 'medium' once per 50 years; for 'high' once per 5 years. The two extreme classes, and two types of uncertainty are used as they are for impact. Note that whereas the likelihood is assessed for one particular threat on one particular node, the impact is assessed on its effect on the telecom service as a whole.

For each node, likelihood and impact are assessed for each threat on that node. Based on these assessments, an overall risk level for that node is calculated; see the RASTER manual for a description of the computation method. Stage 2 is complete when this calculation has been done for every non-actor node.

In the *third stage* (common cause failures analysis), common cause failures (CCFs) are analysed. A CCF means that a single event causes damage to two or more nodes, affecting the availability of the service. For example, a single digging incident can cut two, otherwise independent underground cables. The importance of CCFs in safety and reliability has been recognised for a long time. The initial design analysed CCFs by inspecting each matching pair of nodes (where matching means that the nodes are of the same type, or one of them is an unknown link). One of the major outputs of the case study was a complete revision of this stage; the method now requires the analysts to define clusters, based on geographical proximity of nodes, or a similar critical property. CCFs are then analysed on the level of clusters, instead of nodes. This greatly reduces the effort required, and will be described in more detail in the section *Case study execution*. As described above, assessing single and common cause failures may involve expanding unknown links, making the model more detailed.

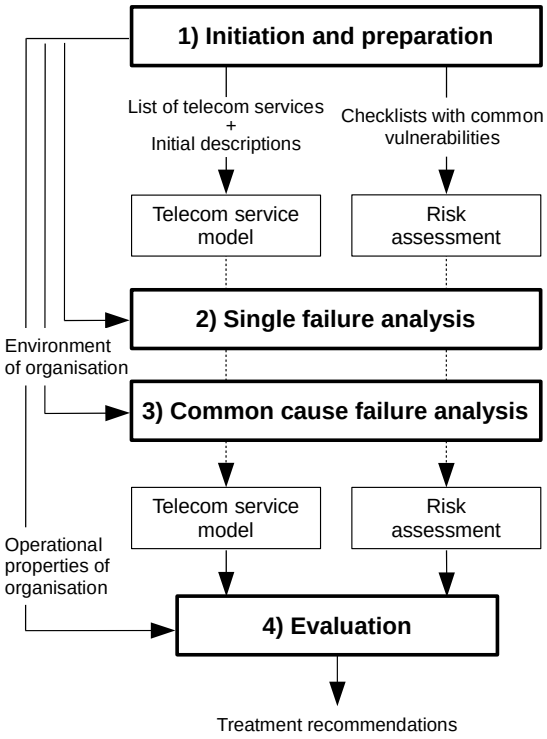

Figure 2: Overview of the method, showing the stag and their inputs and outputs. Stages of the method a shown in bold. Plain arrows denote input and outpu the dotted arrows indicate stepwise refinement.

The *fourth stage* (evaluation) is the final risk evaluation. Based on the analysis in stages 2 and 3, a longlist of risks is compiled. Each risk on this list is the combination of a threat and a diagram node. A prioritisation is necessary, as the decision maker has limited resources and needs to known which risks should be treated first. Not all risks on the longlist are equally important. For example, suppose that the crisis organisation critically depends on some telecom service. This service may be highly reliable, but for additional safety a secondary backup service can be employed. That secondary service can be far less reliable, as long as it does not share failure modes with the primary service. A high risk failure of the secondary service is then far less relevant than a lower risk failure of the primary service. As another example, some services may be "nice to have" but not essential; high risks of failures in those service is therefore less relevant for the effectiveness of the crisis organisations and may in some cases even be ignored. The analysts therefore prioritise the longlist, dropping some risks in the process. The resulting shortlist is the basis for treatment recommendations. Since we observed that decision makers face public scrutiny of their risk management decisions, treatment recommendations must take possible public responses into account. The analysts do so by assessing a fixed list of additional risk factors, which we have called treatment factors and attitudes, to contrast them with the physical factors mentioned earlier (Vriezekolk et al. 2011b). These factors assess the amount of control and oversight, and the personal and shared attitudes to risk of external stakeholders. The resulting risk descriptions and their treatment recommendations are presented to the decision makers, concluding the method.

A full manual of the method is available at the RASTER website (Vriezekolk, 2011c).

Design decisions and requirements

We now describe how the requirements influenced our design decisions. Requirement R1 (No full model) we addressed by the use of unknown links ("cloud shapes") in our diagrams. Using unknown links, entire subsystems can be modelled using a single node, so that not each part of the telecom service has to be modelled in equal detail. Unknown links also help satisfy Requirement R2 (Uncertainty). R2 is further met by the inclusion of likelihood and impact classes for 'plain' uncertainty and for ambiguity. Requirement R3 (Risk factors) is met by assessing physical factors in the analysis of single and common cause failures; treatment

factors and attitudes are assessed in the evaluation stage. We intend requirement R4 (Usability) to be met by the use of a graphical modelling tool. This must be validated by empirical work, such as the case study reported below. The two classes for extremely high and low likelihood and impact are introduced because of requirement R6 (Exceptional circumstances). This requirement is further met by the fact that likely scenarios are identified in stage 1 (initiation and preparation) and used to drive the analysis in stages 2 and 3.

VALIDATION AND CASE STUDY

In this section we describe how the initial RASTER method was validated in practice using a case study, and how this validation resulted in a number of improvements to the method. In section Research Method we formulated three questions. We answered these questions by reviewing the analysis and recommendations created by the participants, and the outcomes of the questionnaire and interview that were conducted after the case study.

Case description

The case study was performed at the institution of the first author: Radiocommunications Agency Netherlands. This is a Dutch government agency responsible for operational activities and enforcement of regulations pertaining to telecommunications. The agency is located in two offices, separated by about 180 kilometres. The internal crisis organisation of the agency forms part of the broader crisis organisation of the Ministry of Economic Affairs, Agriculture and Innovation. Outside crisis response phases, the Agency's crisis team advises on usage and technical aspects of telecom infrastructures. During crisis response, the Agency's crisis team monitors the radio spectrum, locates and resolves interference sources that affect wireless communications, and advises decision makers on technical issues concerning telecom infrastructures.

Five analysts participated in the study. Two of the analysts were active members of the crisis management team; all had experience with technological aspects of telecommunications systems. The responsible managers supported the case study, and allowed the analysts to spend some of their time on the study. At the start of the case study, the analysts found that nine telecom services were being used in crisis response: mobile telephony, fixed telephony to the public telephony network, internal 'voice over IP' (voip) telephony between the two office locations, the national emergency telephony system, email, satellite phones, pagers, office automation via remote desktops, and a video conferencing link between the two offices. To reduce time demands on the analysts, they decided to select two services for further analysis: internal voip telephony, and satellite phones. The analysts chose these two, because they deemed them to be the most complex and the least known.

Case study execution

The analysts held about a dozen two-hour meetings for the case study. The first author was present at most meetings as a facilitator; he was not involved in actual execution of the method. In the first meeting this facilitator introduced the analysts to the RASTER method, and to the object of the case study. The analysts received a written description of the method for reference. From the first introduction onwards, the analysts offered alternatives to the method and possible improvements to it. Their attitude was critical, but supportive. The analysts then collected information for stage 1, the preparation stage. They found that uncovering the required information was easy, as most of the information had already described for their own internal use in a format suitable for the case study. The analysts divided themselves into two teams, one team in each office location. Each team resolved one of the two preselected telecom services.

During stage 2, analysis of single failures, it quickly became apparent that tool support was needed. The analysts tried to create and record the diagrams and the scoring of vulnerabilities using common office automation software. This proved to be cumbersome, as information on diagrams and vulnerabilities of nodes was kept in disconnected applications. When adding or removing diagram nodes, the list of vulnerabilities had to be updated manually.

The case study was therefore paused, and we created a browser-based tool, using standard web technologies and common Javascript libraries. The analysts volunteered in testing the tool, and offered many suggestions for improvement. For a screen shot of this tool see Figure 3.

After the RASTER tool was introduced, the case study restarted at stage 2 (analysis of single failures). During stage 3 (analysis of CCFs), the RASTER tool automatically presented the list of potential CCFs that had to be assessed. From this list, only a small fraction turned out to be realistically vulnerable to a CCF. Out of 351 pairs, the analysts considered only 14 (4%) to be relevant; they concluded that for all other pairs no CCF was

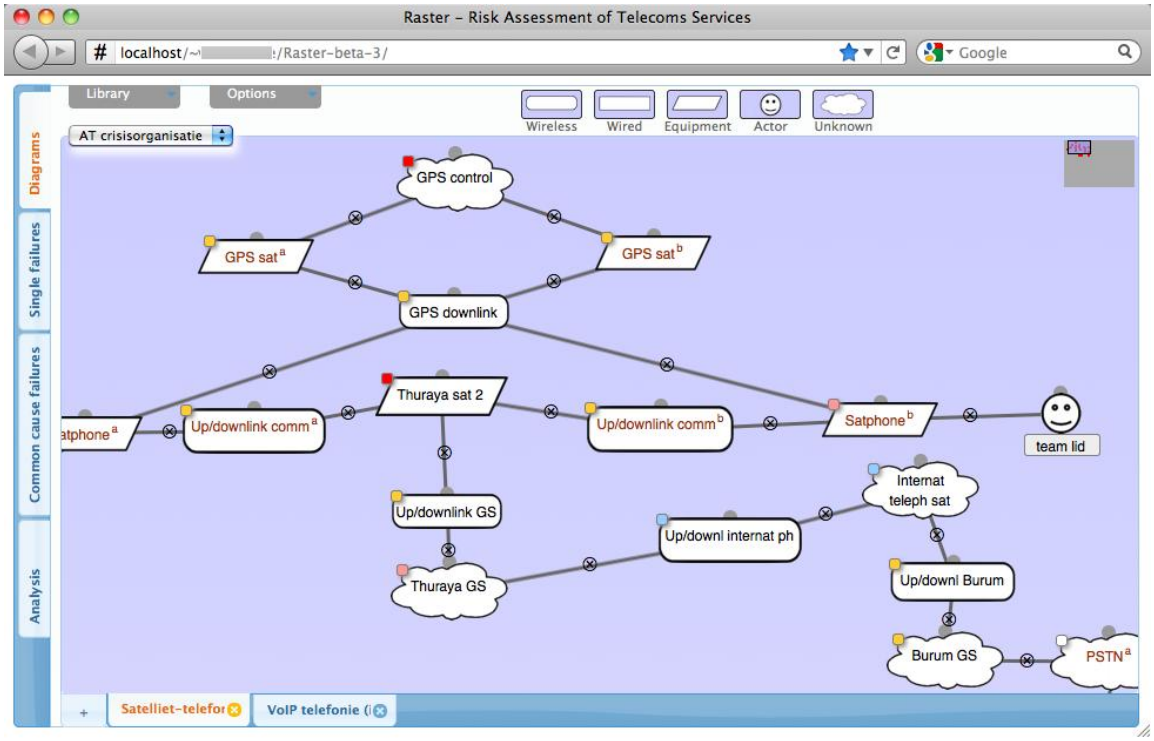

Figure 3: A screen shot of the browser-based tool that was created to support creation of diagrams and the analysis of single and common cause failures. The coloured emblems on the diagram nodes indicate their overall risk level.

reasonably possible. Although the method appeared to yield correct results, the analysts found the procedure too cumbersome and time consuming, thus violating requirements R4 (Usability) and R5 (Effort).

We then looked at alternative approaches for the CCF stage. Many node pairs can be dismissed out of hand because the geographical distance between them is too large to make their common failure probable or possible. We considered adding location information to nodes, as well as a maximum effect-distance to each threat, so that the tool could automatically reject node pairs for which a CCF was impossible due to their distance. For example, a fire would only affect equipment within the same building, whereas flooding would affect equipment in a much larger area. By allocating 'fire' an effect-distance of 50 metres and 'flood' an effect-distance of 30 kilometres, the tool could automatically compute all potential pairs of components for which a CCF is possible. This idea was rejected, because we expected that location information would be difficult to obtain, and that the additional complexity would make the method impractical. Furthermore, not all threats are physical in nature. Configuration errors, for example, can affect two components regardless of their physical separation. Proximity in this case means "being maintained by the same support team". For threats such as Aging, proximity is determined by replacement policies determined by ownership or responsibility.

Instead, we adopted an alternative that was suggested by the analysts themselves. Nodes are grouped into clusters that represent geographical areas, such that CCFs are confined to nodes within each cluster. The analysts expected that they would be able to quickly and easily create such clusters, based on their knowledge of the approximate physical placement of nodes. Instead of a set of geo-coordinates, it would be sufficient to know the region in which the node was to be found. Clusters can be created for other proximity-measures as well, so that threats such as configuration errors can be treated using the same method. Again, we paused the case study while the browser-based tool was updated to reflect this new approach to CCFs.

The case study was continued and completed in a single full-day session, using the latest version of the method and the tool. Stages 1 and 2 could be reused without any modification, so the case restarted at the revised stage 3. The analysts created node clusters, and performed their analysis. For stage 4 (evaluation) the longlist and shortlist were created. Time constraints prevented the analysts from completing stage 4 in full, but they did briefly describe the risk factors for treatment aspects and attitudes for each risk on the shortlist. They also discussed possible risk treatments.

Evaluation and questionnaire

After completion of the case study, each analyst filled out a questionnaire. The purpose of this questionnaire was to obtain data with which we could improve the method for analysis of single and common cause failures. In the first question we asked the analysts for each threat identified in the case study about the influence of each physical risk factor. Possible answers were positive (the factor does influence the risk severity), neutral (factor has little or no impact on the risk severity), or blank ("don't know/unsure"). For each risk factor we calculated the weighted sum of positive and neutral responses; weights were based on the number of times the threat occurred as well as the level of the threat, as assessed in the case study.

For the second question we selected eight node–threat combinations (four from each telecom service). These eight were selected to give a level number of low, medium, and high scores for likelihood and impact. We then asked the analysts to estimate upper and lower bounds to the value of each physical risk factor, for the combinations from the telecom service that they analysed.

Finally, we held semi-structured final interviews with the analysts individually.

RESULTS AND DISCUSSION

We first present some basic data. For the two telecom services combined, the analysts modelled 46 non-actor nodes: 6 wireless links, 13 wired links, 17 equipment items, and 10 unknown links. In the single failures stage 176 node–threat combinations were analysed. Of these, 39 were ranked extremely low, 83 as low, 30 as medium, 7 as high, 3 as extremely high, 14 as unknown, and none as ambiguous. In total, the analysts created 41 node clusters, for 17 threat types. Hence, in the common cause failures stage 41 node clusters were analysed. For stages 2 and 3, a total of 217 threat assessments were completed. The checklist with common threats contained 7 threats to equipment items, 4 to wired links, and 4 to wireless links. In addition to these 15 default threats, the analysts added 7 other threats in the single failures stage.

During the evaluation stage, a longlist of 26 risks was compiled: 21 single failures and 5 common cause failures. Of the single failures, 13 were ranked as unknown, 5 as high, and 4 as extremely high. Of the common cause failures, 3 were ranked as high, and 2 as medium.

The analysts estimated to have spent 40 hours on the entire case study, and about half that time on executing the RASTER method itself. Anonymised case study and interview results can be requested from the first author.

The analysts did not rank any threat as ambiguous. They viewed ambiguous scores mostly as a failure of discussion of the underlying reasons for scores. This does not necessarily mean that ambiguity is not a useful concept, but it does suggest that further research may be necessary. Although 71% of all threat types originated from the checklists, the checklists accounted for 95% of all analysed node–threat combinations. Added threats were typically used for a single node only. From this we tentatively conclude that checklists are a suitable method for threat identification and help save time.

We also observed that the analysts seldom referred back to the descriptions of the seven likelihood and impact classes after having read the manual. This was also reflected in the outcome of the questionnaire; there was little correlation between the frequency-score assessed in the case study, and the risk factor likelihood as estimated in the questionnaire. Similarly, there was little correlation between the impact-score assessed in the case study and the estimates for magnitude and geographical and temporal extent of the damage or repairability. To verify sensitivity, we replaced the frequency and impact scores in the actual case study with values based on the estimates given in the questionnaires. In all but one case this led to a higher overall threat level for that node and, using the line of reasoning the analysts employed, to more risks on the longlist, and possibly the shortlist as well. Accuracy and usefulness of the end results were not goals of this case study. However, in future case studies such imprecision would be more serious, as it leads to inaccurate results when these assessments are compared in the evaluation stage. The interviews also made clear that the analysts require more guidance in the assignment of classes to frequency and impact.

A related observation is that the analysts did not strive for firm justifications of all threat assessments. Their scores were often based on general experience with telecom hardware, not with the properties of specific nodes. For example, the analysts assessed the likelihood of breaks in Ethernet cables as 'low' without searching for specific local historical data to back up this claim. On the one hand this shows a benefit of Raster, in that it does not require all input data to be equally accurate, but as described above in some cases it may have a significant effect. We guess that with a larger group there would have been more interaction, that estimates would have been more accurate, and that missing data would have been flagged sooner.

Interestingly, the RASTER tool showed at a glance that the internal voip telephony service was rated as less risky than the satellite telephony service. Although part of this difference can be explained by differences between the two services (voip telephony may well have higher availability than satellite telephony), the analysts did acknowledge the possibility of an uneven bias between the two self-assigned teams.

The questionnaires confirmed our assumption that the physical risk factor 'delay of effects' is hardly relevant in telecom service availability risks (unlike perhaps in health risks of food additives). Questionnaire results unsurprisingly indicated that the physical risk factors likelihood, extent of damage, and repairability are highly relevant, but did not conclusively show the relevance of geographical and temporal extent of damage. For crisis organisations, the difference between failures with a long repair time and unrepairable failures is not relevant, since crises are typically short-lived. The opinions of analysts on the time required for execution of the method ranged from "relatively little" to "quite a bit". Overall they did not expect to need less time for future studies, as they expect to spend most time on discovering the components and their interconnections.

Finally, we return to our three questions to determine applicability of RASTER, as listed in the Research Method section. With tool support and with the latest version of the RASTER method, the analysts were able to use and complete the method (1). Although time constraints prevented the completion of stage 4 (evaluation) in detail, they were able to derive a shortlist and discuss suitable treatment recommendations (2). Based on the final interviews, the analysts were mostly satisfied with the results. The analysts found the overall amount of time reasonable, but did not expect to need less time in future (3). We believe that this is an important result, and the main finding of this case study.

LESSONS LEARNED AND FURTHER WORK

The case study gave us valuable starting points for three practical guidelines for applying the RASTER method: group size (i), division of tasks (ii), and method organisation (iii).

(i) A group of five analysts participated in this case study. Since the analysts decided to split themselves into two groups, each group addressing one telecom service, the actual number of analysts involved in analysing each service was less. For purposes of validating the method, having five analysts was sufficient. However, for a full application of the method it would be beneficial if a larger group could be employed, in order to enlarge the available skill set and to promote discussion. The optimal group size probably depends on the organisation, the distribution of skills, and the complexity of the telecom services. For further case studies, we would prefer to have at least 8 analysts participating.

(ii) We are unsure whether it is beneficial to split the group into sub-groups, each analysing one telecom service. Although this may appear to be more time-efficient for stage 2 (single failures), we have two reasons to think that it makes other stages less efficient. First, stage 3 (common cause failures) is done for the project as a whole, not for each individual service; each analyst would therefore need to understand each telecom service in order to participate in the analysis of common cause failures. Secondly, a consistent estimate of frequency and impact is required in stage 4 (evaluation). With separate groups analysing each service, it will be more likely that an uneven bias leads to inaccurate evaluation results. However, creation of the initial diagrams could perhaps be done by different subgroups in parallel.

(iii) The manuals and organisation of RASTER need further thought. More guidance is needed on the scoring system for frequency and impact of threats. A meeting secretary or facilitator can help ensure the quality of the assessments and can help maintain consistency in scoring.

In conclusion, we have developed the first version of a usable RA method for telecom service availability risks that takes into account the special circumstances in which crisis organisations operate. We have shown that our method can deliver results in a limited but nevertheless real-life situation. To our knowledge, no other RA method can do so with the same ease and speed. We plan at least two further case studies. One will test whether RASTER is robust against change of the team of analysts, and of their level of expertise; it will have independent teams executing the method on a common case. The other will test RASTER's performance for a complex public crisis organisation.

ACKNOWLEDGMENTS

We wish to thank the five analysts from Radiocommunications Agency Netherlands for their invaluable contribution, and for their many suggestions for improvement of the RASTER method.

REFERENCES

1. Barber, B. and Davey, J. (1992) The use of the CCTA risk analysis and management methodology CRAMM, *Proc. MEDINFO92, North Holland,*
2. Fredriksen, R., Kristiansen, M., Gran, B., Stølen, K., Opperud, T., and Dimitrakos, T. (2002) The CORAS framework for a model-based risk management process, *Proceedings of the 21st International Conference on Computer Safety, Reliability and Security,* 39–53.
3. Hale, A. R. and Hovden, J. (1998) Management and culture: the third age of safety. A review of approaches to organizational aspects of safety, health and environment, *Occupational injury: Risk, prevention and intervention,* 129–165.
4. Hollnagel, E. and Goteman, O. (2004) The functional resonance accident model, *Cognitive System Engineering in Process Plant 2004*.
5. ISO (2009) Risk management – Principles and guidelines, International Standard 31000:2009.
6. Johnson, W. G. (1973) Management Oversight and Risk Tree – MORT, *Aerojet Nuclear Co., Scoville, ID (USA).*
7. LaPorte, T. R. and Consolini, P. M. (1991) Working in Practice but Not in Theory: Theoretical Challenges of "High-Reliability Organization", *Journal of Public Administration Research and Theory: J-PART,* 1, 1, 19–48.
8. Leveson, N., Dulac, N., Marais, K., and Carroll, J. (February/March 2009) Moving Beyond Normal Accidents and High Reliability Organizations: A Systems Approach to Safety in Complex Systems, *Organization Studies,* 30, 2–3, 227–249.
9. Perrow, C. (1984) Normal accidents: Living with high-risk technologies, Princeton Univ Pr.
10. Radiocommunications Agency Netherlands (2009) State of the Ether ("Staat van de Ether", in Dutch), *Agentschap Telecom (http://www.agentschap-telecom.nl).*
11. Radiocommunications Agency Netherlands (2010) State of the Ether ("Staat van de Ether", in Dutch), *Agentschap Telecom (http://www.agentschap-telecom.nl).*
12. Reason, J., Shotton, R., Wagenaar, W., Hudson, P., and Groeneweg, J. (1989) TRIPOD, A Principled Basis for Safer Operations, *Report prepared for Shell Internationale Petroleum Maatschappij, Exploration and Production.*
13. Swain, A. D. and Guttmann, H. E. (1983) Handbook of human-reliability analysis with emphasis on nuclear power plant applications. Final report, *Sandia National Labs, Albuquerque, NM (USA).*
14. Vriezekolk, E., Wieringa, R., and Etalle, S. (2011a) A new method to assess telecom service availability risks, *Proceedings of the 8th International Conference on Information Systems for Crisis Response and Management ISCRAM2011*.
15. Vriezekolk, E., Wieringa, R., and Etalle, S. (2011b) How to assess telecom service availability risks for crisis organisations?, *Advances in Safety, Reliability and Risk Management,* 2653–2661, London.
16. Vriezekolk, E., (2011c) Raster documentation, http://wwwhome.ewi.utwente.nl/~vriezekolke/Raster/ .
17. Wieringa, R. J. (2009) Design Science as Nested Problem Solving, *Proceedings of the 4th International Conference on Design Science Research in Information Systems and Technology, Philadelphia,* 1–12, New York.
18. Williams, J. C. (1988) A data-based method for assessing and reducing human error to improve operational performance, *Fourth Conference on Human Factors and Power Plants, 1988, Conference Record,* 436–450.