Mapping the Information Flows for the Architecture of a Nationwide Situation Awareness System

(Poster)

Hayretdin Bahşi
Tallinn University of Technology
Tallinn, Estonia
hayretdin.bahsi@taltech.ee

Veiko Dieves
Estonian National Defence College
Tartu, Estonia
veiko.dieves@mil.ee

Taivo Kangilaski
Tallinn University of Technology
Tallinn, Estonia
taivo.kangilaski@energia.ee

Peeter Laud
Cybernetica AS
Tartu, Estonia
peeter.laud@cyber.ee

Leo Mõtus
Tallinn University of Technology
Tallinn, Estonia
leo.motus@taltech.ee

Jaan Murumets
Estonian National Defence College
Tartu, Estonia
jaan.murumets@mil.ee

Illimar Ploom
Estonian National Defence College
Tartu, Estonia
illimar.ploom@gmail.com

Jaan Priisalu
Tallinn University of Technology
Tallinn, Estonia
jaan.priisalu@taltech.ee

Mari Seeba
Cybernetica AS
Tartu, Estonia
mari.seeba@cyber.ee

Ermo Täks
Tallinn University of Technology
Tallinn, Estonia
ermo.taks@taltech.ee

Kaide Tammel
Estonian National Defence College
Tartu, Estonia
kaide.tammel@mil.ee

Piia Tammpuu
University of Tartu
Tartu, Estonia
piia.tammpuu@ut.ee

Kuldar Taveter
Tallinn University of Technology
Tallinn, Estonia
kuldar.taveter@taltech.ee

Avo Trumm
University of Tartu
Tartu, Estonia
avo.trumm@ut.ee

Tiia-Triin Truusa
University of Tartu
Tartu, Estonia
tiia-triin.truusa@ut.ee

Triin Vihalemm
University of Tartu
Tartu, Estonia
triin.vihalemm@ut.ee

***Abstract*—A nation-level situation awareness (SA) system will facilitate the making of right decisions by serving the decision-makers with the information they need. In this paper, we show how the mapping of information flows between governmental institutions allows us to develop such a SA system and continuously improve it by optimizing the movements of information. To this end, we propose a systematization method for the various kinds of information flows, and explain how it helps to reduce their perceived heterogeneity.**

***Keywords*—situation awareness, comprehensive national defense, system architecture**

I. Introduction

Situation awareness (SA) means knowing and understanding what is going on and what is going to happen. Depending on the complexity of application, the procedures for assessing SA and methods for SA distribution to decision-makers may rely on different approaches, e.g. [1]–[3], to better satisfy the requirements of information owners and to ensure reaching the goals set by stakeholders A nation-level SA system supports the decision-makers at the various locations and levels of the government, the local authorities, as well as actors from the private sector, providing them with the information that enables them to improve the overall functioning of the country. Such a system will allow the threats to be noticed and handled in an early stage, either before they have had the chance to cause real harm to the nation, or at least when the harm they're causing is minimal.

The heterogeneity and the number of different information flows in such a SA system is enormous, hence there are no publicly known working examples of nation-level SA systems. Detection and assessment of interdependencies between those information flows is even more challenging. Recently, Estonia has ventured to build a "comprehensive system for nation-level SA, supporting the national defense in the broad sense" [4]. Here the "broad sense of national defense" concerns the activities of not only the Ministry of Defense and the Defense Forces, but also the activities aimed to prevent and manage of natural or man-caused disasters, and/or transforming one into another [5]. Estonia's small size and its excellent application of ICT in the public sector are important factors that promote Estonian efforts for building a prototype of a nation-level SA system [6].

In this paper, we discuss the first stage of designing the architecture for a nation-level SA system by fixing the existing and/or required information flows. The future papers will discuss second stage of specifying the architecture for SA system and will be concerned with building models for the institutions generating/consuming information flows, and with analyzing

This research has been funded from the European Regional Development Fund through the Strengthening of Sectoral R&D (RITA) activity of the Estonian Research Council (ETAg). ETAg has also supported this research through grant IUT27-1.

the use of existing in-institution ICT systems. Also part of the second stage elaborates the methods for checking spatial and temporal consistency of the provided information. The third and final stage of building the architecture for prototype SA, focuses on information flows enabling interoperation of nation-level institutions, providing decision-makers on different levels of the system with consistent information, technical sustainability of communication network, and capability to simulate the impact of some key decisions. We will discuss these steps in more detail in Sec. III.

We discuss the discovery of existing and required information flows in Sec. IV. By considering the existing flows of information, we will learn the available sources of information, how the information is processed and consumed. We can decide what kinds of resources are needed for the processing and transmission of information while preserving the non-functional requirements of information handling. We can then optimize the existing operations, altering the paths that certain pieces of information take, with the aim of improving the precision or the resiliency of the information available for decisions, or even removing redundant pieces of information.

This paper makes the following contributions:

- It identifies the goals and non-goals of a nation-level SA system, based on the needs of its users and the generalities it has to deal with.
- It puts forth a systemic manner for describing the information flows among the governmental departments, as well as private organizations that participate in comprehensive national defense.
- It proposes the steps for collecting the descriptions of information flows, as well as steps that allow the collection of these descriptions to be developed into the SA system architecture.

II. Goals and challenges of a nation-level SA system

A decision-maker, either in the government or in the private sector, follows the *key performance indicators* (KPIs) of the systems he/she is responsible for, as well as the *key risk indicators* (KRIs) of their environment to arrive at satisfactory decisions. We see the main role of a nation-level SA system to be the collection, movement, processing, and visualization of the information about the current state of the country, and presenting it to the decision-makers.

A nation-level SA system is actually the SA system for a large System of Systems (SoS). In Estonian setting, we can estimate its complexity by considering the number of institutions and enterprises exchanging information with each other over the X-road infrastructure. X-road [7] is the data exchange layer used by all Estonian public institutions, as well as a number of private utility companies (telecom, energy, banking). At the end of 2018, X-road facilitates the data exchange between more than 650 institutions and enterprises[1]. The Estonian nation-level SA system would cover a similar number of systems. This SA system has to detect majority of potential threats against this SoS in due time, so that they can be avoided, or their unwanted effect on country's functioning can be mitigated.

The relevant flows of information present in this SoS are extremely varied in terms of the content and the kind of information, the bandwidth necessary, as well as the timing and security properties. This heterogeneity is the main challenge in designing the system, we need to tackle it both in the design phase, and when proposing the actual architecture.

Due to this heterogeneity, we cannot see the tasks of a nation-level SA system to go much beyond ensuring that the right information is at the right place in the right time. While SA systems for more homogeneous systems readily accommodate higher-level information fusion and automated drawing of conclusions [8], a nation-level SA system can only support decision-makers at various institutions applying their system-specific experience. Similarly, we expect there to be no uniform method to visualize the incoming information flows. Instead, it will happen at each institution and at each decision-making level in a manner that is most suitable to this place.

On the other hand, we expect a nation-level SA system to help optimize the information flows in a nation. The existing flows have become established in ad-hoc manner, with some local optimality in mind. The analysis for the SA system will give us a clearer view of the information available in, and of use in various institutions. Using this view, we can make well-informed suggestions, which information flows there should be between institutions. A well-engineered mechanism for setting up information exchanges allows us to deploy the suggested changes.

With the goals set like this, a nation-level SA system promotes the achieving of *distributed situation awareness* [9] in regards of the events and vulnerabilities that are significant for national defense. In this way, each institution and each decision-maker is well aware of the details they use in their own decision-making, but there does not have to be a single location (although our proposed architecture does not exclude the creation of such location) where the entirety of the information is collected and fused.

III. Proposed methodology

We propose to perform the system analysis for a nation-level SA system through the following steps. The system analysis leads to an understanding of the quantitative and qualitative needs of information processing, which leads to the architectural details of the SA system. The steps named here may be repeatedly fulfilled, leading to a more detailed view of information flows, and a more informed deployment of the SA system.

- Identification of nodes that create, store, process or consume information. These nodes may be institutions, or their departments, or certain positions in them, or certain information systems. These are derived from the normative documents of the institutions, as well as from the interviews.

[1] https://www.x-tee.ee/factsheets/EE/#eng

- Describe the existing information flows, filling in the Pre-structured Table of Information Collection (PTIC). Similarly to the previous point, these are collected from normative documents and from interviews. An additional source for characterizing inter-institutional information flows are the various logs of existing systems. We give the details in Sec. IV. The resulting table can be seen as the consolidated observation report of interactions between the relevant institutions, where these interactions somehow contribute to the decision-making. Besides the interactions, the interviews will also result in descriptions of decision-making processes (as a textual narrative). These descriptions may help in performing the next steps in our methodology.
- Identify the key resources that the information flows report on, and that the decision-makers use to evaluate the KPIs and KRIs. These are obtained from the PTIC, as well as from the descriptions of decision-making processes.
- Build models of institutions. These models allow to identify the escalation and de-escalation points, where this institution (as a system) starts to behave differently. These models are derived from the KPIs, as well as from the information flows in the PTIC. The models allow us to pin-point the information needs of the institutions.
- Propose the architecture of the SA system, identifying the computational nodes of this system (likely to largely coincide with the information creation, storage, processing, and consuming nodes identified in the first step), their information processing and storage needs, as well as the necessary parameters of information transmission between nodes (bandwidth, latency, security, etc.). We intend to leverage an existing technical architecture for the computational nodes as much as possible, this architecture is described in Sec. VI.

We describe the first two steps of this methodology in more detail in the next Section. The following steps will be discussed in Sec. V and detailed in future work.

The identified information flows can be used for other purposes beside the fixing of the system architecture for nation-level SA. We can use them to look for weaknesses in the current governance models. We can also use them to identify dependencies between the different systems that are significant from the comprehensive national defense point of view. Both of these analyses give valuable input to the optimization of information flows between institutions. These analyses will be detailed in future work.

IV. Information collection and ordered representation

In order to collect and systematize the information flows, we need a taxonomy of the heterogeneous functional information needs of institutions contributing to comprehensive national defense (CND). This taxonomy is a methodological solution for mapping the necessary information and its flow between different institutions. It distinguishes between different management levels (tactical, operational, and strategic) and turning points in crisis management (change of leadership; escalation).

We have materialized this taxonomy as the *pre-structured table for information collection* (PTIC, Table I). It characterizes the source and the target of a particular flow of information, the functional characteristics of this flow (e.g. the type of transmitted data, its bandwidth, burstiness), as well as its non-functional requirements, including security and timeliness. PTIC also describes the semantics of the transmitted information, telling us what kind of information is considered to be important.

Let us discuss the meaning of some of the columns of the PTIC. While the goals of the actor are stated in language specific to the statutes applying to that actor, the task and subtask are stated in actor-independent manner. For this we need a standardized list of tasks. While filling the PTIC, we also need to come up with this list. We intend to start with the Mission Essential Task List of NATO, as this is a rather comprehensive existing list that covers a multitude of activities pertaining to the preparation, handling, and recovering from a crisis. We expect to expand that list during the collection of information flows.

The "data owner" and "needed information" specify the source of the information as precisely as possible. We aim to name the information system that serves as the source. In Estonia, the entries in the catalog of the state information system's management system[2] can be used.

The "type" column describes whether this information flow is generated by the agent, or it is received from some other actor. It also indicates whether it is directly consumed, or receives some kind of processing (e.g. aggregation) at its source. Note that a possible type of information are the decisions made at other institutions.

The data format and frequency characterize the functional requirements on information transmission channels. The "security class" column describes the non-functional requirements. Here the integrity and availability requirements are specified by the recipient of the information, while the confidentiality requirements are specified by the sender. We will use a standardized list for describing the desired security levels.

In general, we aim to obtain standardized descriptions for as many aspects of the description of an information flow as possible. All this is done for the purpose of reducing perceived heterogeneity, and put into use in the following steps of coming up with the SA system architecture.

As the aim of the nation-level SA system is to contribute to comprehensive national defense, the collection of data begins with identifying the tasks that form part of comprehensive defense. Using a standardized mission essential tasklist, we detect the tasks relevant for comprehensive national defense. From documents (mainly laws and statutes) we map the goals and objectives of agents /organizations that carry out these tasks. The PTIC gives an overview of the tasks that an agent

[2]https://www.riha.ee

Name	description	possible values
Agent	Actor (an institution, or a decision-maker at a specific management level) that performs a specific task	Name of an organization or its structural division
Goals, as stated	Goals/objectives, as stated in normative documents (laws and statutes)	Name of the normative document, and a fragment from it
Cooperations	A list of other actors, with whom information is shared when tasks are performed, and their role in a specific task.	Names of the partnering organizations or their structural divisions
Task & subtask	A comprehensive national defense related task that relates to the goals and objectives of an agent	An element from a standardized list of possible tasks
Data owner	Organizational entity that is responsible for storing, securing and sharing the data	Name of the organizational entity and/or an information system that stores the data
Needed information	The phenomenon about which the information is requested	"situation", "environment", "impact", "estimated time", etc.
Type	Is this primary information (collected by the agent) or the result of some analysis?	"primary", "analysis"
Format	The data type	Number, text, image, video, database (with given schema), etc.
Frequency	How often and how regularly is the information transmitted?	"once", "regularly" (incl. frequency), "ad-hoc" (either "push" or "pull")
Data source	In case of primary data: how is this piece of information collected?	"observed by a human", "collected by sensors", etc.
Security class of data	The security class consists of three components, describing the *confidentiality*, *integrity* and *availability* requirements for the data.	Each of the three components comes from a standardized list.
Data channel	Description of the channel that is used to transmit information	"well-defined API", "informal discussion between officials face-to-face / over phone", etc.
Data validity period	For how long can the receiving party rely on the received information?	a time interval, and/or the description of an event (e.g. "until the next transmission")
Data validity region	To which geographic region does the information apply to?	town, county, state, operational areas, etc.

TABLE I
COLUMNS OF THE PRE-STRUCTURED TABLE OF INFORMATION COLLECTION

performs for comprehensive national defense, its main partners with whom the information is shared and information about the data that is being shared.

PTIC is filled in for each relevant agent/organization using document analysis, or exchange of information with the representatives of agents, or data mining techniques applied to the communication or transaction data. These approaches can validate or complement the findings identified in each other as one approach may not provide the whole picture or its findings may not reflect the actual business processes. The experts may not know all the necessary details or they may not aware the recent changes in these processes. On the other side, although data mining techniques have been developed for extracting the details of business processes from the transactional or event data [10], the results of these techniques cannot cover the whole problem domain as they are limited to the type and amount of input data and may not reveal all aspects of context knowledge.

In Estonian setting, data mining techniques can be applied to the transaction data on X-Road services, official documents exchanged through Document Exchange Center or official meeting protocols in order to derive the communication patterns of the institutions. The work-related phone calls and e-mails of individuals with official responsibilities can act as another data source to enrich this pattern derivation task.

A number of interviews will be conducted to fill the PTIC. These interviews will fall into three major classes. First, we interview the senior specialists in institutions, learning from them the possible roles and activities present in the institution, and the information that is used by them. Second, we interview the members of ad-hoc crisis management committees, learning from them the actual information needs they had in while handling the crisis. Third, we consult the technical personnel in order to fix the data format and other technical aspects of the transmitted information.

The filled PTIC will be validated through subsequent interviews, as well as through a war-game.

V. NEXT STEPS

After filling the PTIC, we will use it to reduce the perceived heterogeneity of the SoS for which we are creating the SA system. A major next step is generating the models of institutions that represent information flows, events, decision points and decision processes between analyzed organizations. We create an organization model of the institutions responsible for ensuring situation awareness and reacting to undesired events. The organization model represents the organizational roles and relationships and dependencies between the roles, such as control, peer, and benevolence relationships [11], as well as goal, task, and resource dependencies [12]. This is followed by modeling each organizational role in terms of its responsibilities (i.e., what an agent performing that role should do) and constraints (i.e., what an agent performing that role should not do, if any) [13]. As the next step, information flows between the organizational roles are modeled as information flow diagrams [14] based on resource dependencies between the roles. Finally, the information flows with decision points and decision-making processes are represented as processes spanning different organizational units. The variety of models for the same institution is necessary in order to detect the inconsistencies in the information flows [15]. From these models, we can elicit the technical requirements for the SA system and proceed with developing the system architecture.

We can use the models and the outputs of the war-games used to validate the PTIC to perform a vulnerability analysis for the entire SoS. We will have modeled the dependencies and relationships among assets, services or processes, and also models how assets are used by services and services invoked by processes. Such impact dependence graph [16] shows how the failure or reduced availability of a certain component affects other components reachable from it through one or more steps. The graph may be annotated by specificities of dependencies (e.g. they may vary over time) and the analysis of the graph can take it into account. For performing the vulnerability analysis, the impact scale estimation standard is developed and the impact of previously analyzed cases is mapped to this standard. For each subsystem, direct impact and maximum cumulative effect in case of failure is estimated. Alternative routes are counted and performance degradation is estimated. A list of choke points with no alternative routes is generated. Integrity of each subsystem is analyzed by estimating maximum impact of each channel information modification. If applicable, an alternative route for the same information is described. In the last step, the most dangerous availability and integrity vulnerabilities are combined and most dangerous combinations are found.

The results of the vulnerability analysis will form a part of the basis for proposing improvements for the information flows in the SoS. We expect to use the methods of knowledge management [17] while forming these proposals; they will stem from our suggestions to improve processes inside different organizations. As the output, we will create a table similar to PTIC, but with the information flows we suggest to have. These suggestions can again be validated through similar modeling, war-games, or analysis.

VI. The architecture and the deployment of the SA system

When discussing the architecture of a nation-level SA system, we can distinguish two aspects of it. One is the *system architecture* — where the components of the SA system are located, how powerful are the components, how are they connected to each other, etc. The other aspect is the *technical architecture* — what components exist, how are they built, how can they be configured, etc. The discussion in this paper has been about fixing the system architecture of a SA system, as this depends on the institutions we have in the nation and on the particularities of the information exchanges between them. On the other hand, we have a quite good idea about the technical architecture, for the following reason.

The SA system has to transmit information between different locations and different governmental institutions. It should have no single points of failure. In particular, if the network connection is down in some geographic area or in some institution, then the flows between other areas and institutions should commence unhindered. In Estonia, a system with such property, the X-road, has been deployed for 17 years [18]; it is used for the exchange of information between hundreds of departments of governmental institutions, as well as private companies.

The main information exchanges in X-road take place in peer-to-peer manner, with the X-road security server in one institution directly communicating with the security server in another institution. Behind the security servers, the information systems decide what queries to make and how to answer them. The set of allowed queries — their semantics and representation have previously been agreed upon by the two institutions. The agreement, which is enforced by the security servers, provides an effective method of access control and eases the adoption of X-road usage.

Being a peer-to-peer system, no central component of X-road has to be available for the communication to take place. There do exist certain central components for maintaining the registry of connected information systems and the services they offer, as well for monitoring and maintaining the system. Still, these create no single point of failure. The information exchange is secured, ensuring its confidentiality and integrity. The essential components of the X-road technology may be duplicated, thus increasing the availability of the system. Since its initial deployment, X-road has had zero downtime.

We believe that it makes sense to replicate the architecture of X-road for the SA system, because of the experience gained in running such a system, and on the perceived benefits of the access control mechanisms that the bilateral information exchange agreements bring. This technical architecture has the flexibility to direct the information flows in the manner we have deemed the most suitable, and it will be easy to amend and redirect these flows as necessary.

As X-road has no central components required for information exchange between institutions, the existing configuration, maintenance and monitoring components of X-road may be shared by the SA system, once the latter is deployed. In this manner, we still effectively have a single X-road system, now used to exchange both the regular governance-related information, as well as the information for situation awareness.

References

[1] M. R. Endsley, "Toward a theory of situation awareness in dynamic systems," *Human Factors*, vol. 37, no. 1, pp. 32–64, 1995.

[2] ——, "Situation awareness misconceptions and misunderstandings," *Journal of Cognitive Engineering and Decision Making*, vol. 9, no. 1, pp. 4–32, 2015.

[3] N. A. Stanton, P. M. Salmon, G. H. Walker, E. Salas, and P. A. Hancock, "State-of-science: situation awareness in individuals, teams and systems," *Ergonomics*, vol. 60, no. 4, pp. 449–466, 2017.

[4] "Tervikliku olukorrateadlikkuse võime loomine riigikaitseks (creating a comprehensive situation awareness capability for national defense)," 2018, project call. https://www.etag.ee/wp-content/uploads/2017/12/Olukorrateadlikkus_uus-konkurss_loplik.pdf.

[5] "Eesti julgeolekupoliitika alused (principles of estonian security policy)," 2017, adopted by Estonian Parliament in May 31st, 2017. https://www.riigiteataja.ee/akt/306062017002.

[6] "UN E-Government Survey," 2018, https://publicadministration.un.org/egovkb/en-us/Reports/UN-E-Government-Survey-2018.

[7] A. Ansper, A. Buldas, M. Freudenthal, and J. Willemson, "Scalable and efficient PKI for inter-organizational communication," in *19th Annual Computer Security Applications Conference (ACSAC 2003), 8-12 December 2003, Las Vegas, NV, USA*. IEEE Computer Society, 2003, pp. 308–318. [Online]. Available: https://doi.org/10.1109/CSAC.2003.1254335

[8] F. Baader, A. Bauer, P. Baumgartner, A. Cregan, A. Gabaldon, K. Ji, K. Lee, D. Rajaratnam, and R. Schwitter, "A novel architecture for situation awareness systems," in *Automated Reasoning with Analytic Tableaux and Related Methods, 18th International Conference, TABLEAUX 2009, Oslo, Norway, July 6-10, 2009. Proceedings*, ser. Lecture Notes in Computer Science, M. Giese and A. Waaler, Eds., vol. 5607. Springer, 2009, pp. 77–92. [Online]. Available: https://doi.org/10.1007/978-3-642-02716-1_7

[9] N. A. Stanton, "Distributed situation awareness," *Theoretical Issues in Ergonomics Science*, vol. 17, no. 1, pp. 1–7, 2015.

[10] W. Van Der Aalst, "Process mining: Overview and opportunities," *ACM Transactions on Management Information Systems (TMIS)*, vol. 3, no. 2, p. 7, 2012.

[11] F. Zambonelli, N. R. Jennings, and M. Wooldridge, "Organizational abstractions for the analysis and design of multiagent systems," in *Agent-Oriented Software Engineering, First International Workshop, AOSE 2000, Revised papers*, ser. Lecture Notes in Computer Science, vol. 1957. Springer, 2000, pp. 235–251.

[12] P. Bresciani, A. Perini, P. Giorgini, F. Giunchiglia, and J. Mylopoulos, "Tropos: An agent-oriented software development methodology," *Autonomous Agents and Multiagent Systems*, vol. 8, no. 3, pp. 203–236, 2004.

[13] L. Sterling and K. Taveter, *The Art of Agent-Oriented Modelling*. MIT Press, 2009.

[14] C. Durugbo, A. Tiwari, and J. R. Alcock, "Modelling information flow for organisations: A review of approaches and future challenges," *International Journal of Information Management*, vol. 33, no. 3, pp. 597–610, 2013.

[15] L. Mõtus, K. Taveter, and V. Dieves, "Modelling complex system-of-systems for creating situation awareness : (late breaking report)," in *2018 IEEE Conference on Cognitive and Computational Aspects of Situation Management (CogSIMA)*, June 2018.

[16] G. Jakobson, "Mission cyber security situation assessment using impact dependency graphs," in *Proceedings of the 14th International Conference on Information Fusion (FUSION)*, 2011, pp. 1–8.

[17] J. Girard and J. Girard, "Defining knowledge management: Toward an applied compendium," *Online Journal of Applied Knowledge Management*, vol. 3, no. 1, pp. 1–20, 2015.

[18] A. Ansper, "Designing a governmental backbone," in *Information Security Technology for Applications - 16th Nordic Conference on Secure IT Systems, NordSec 2011, Tallinn, Estonia, October 26-28, 2011, Revised Selected Papers*, ser. Lecture Notes in Computer Science, P. Laud, Ed., vol. 7161. Springer, 2011, pp. 1–3. [Online]. Available: https://doi.org/10.1007/978-3-642-29615-4_1