Trust Strategies and Policies in Complex Socio-technical Safety-Critical Domains: An Analysis of the Air Traffic Management Domain

Massimo Felici

School of Informatics, The University of Edinburgh
Edinburgh EH9 3JZ, UK
mfelici@inf.ed.ac.uk
http://homepages.inf.ed.ac.uk/mfelici/

Abstract. The future development of Air Traffic Management (ATM), set by the ATM 2000+ Strategy, involves a structural revision of ATM processes, a new ATM concept and a system approach for the ATM network. This requires ATM services to go through significant structural, operational and cultural changes that will contribute towards the ATM 2000+ Strategy. Moreover, from a technology viewpoint, future ATM services will employ new systems forming the emergent ATM architecture underlying and supporting the European Commission's Single European Sky Initiative. Introducing safety relevant systems in ATM contexts requires us to understand the risk involved in order to mitigate the impact of possible failures. This paper is concerned with trust in technology. Technology innovation supports further (e.g., safety or performance) improvements, although there is often a lack of trust in changes. This paper argues that organizations need to identify trust strategies and policies supporting the delivery of technology innovation. Moreover, the identification of trust strategies and policies supports the understanding of subtle interactions between diverse, often competing, system objectives.

1 Introduction

Computer systems support diverse human activities (e.g., monitoring, decision making, etc.). The introduction of new computer systems, or the upgrade of existing ones, in any environment often modifies work practice. For instance, system operators often need to adjust their procedures around new systems. Moreover, systems may act as a means of communication/mediation between human beings. Complex interactions [17] emerge as results of changes (e.g., environmental changes, new computer systems, adjusted work practices, etc.). The introduction of new technology often requires the re-negotiation of social organizations (e.g., responsibility and accountability) as well as overall system features (e.g., safety). Change gives rise to uncertainties with respect to computer systems. For instance, in the Air Traffic Management (ATM) domain, air traffic controllers often react to system changes or failures by managing less traffic in their air spaces. Uncertainties require of us an extent of *trust* (e.g., with respect

N. Guelfi and D. Buchs (Eds.): RISE 2006, LNCS 4401, pp. 51–65, 2007.

to computer systems). Unfortunately, changes often trigger mistrust. Norman, for instance, reports how the introduction of questioning between pilots in work practice, initially, triggered a lack of trust in the commercial aviation community [46]. However, the new practice, eventually, produced increased safety[1]. Similarly, empirical studies point out the relationship between trust in automation and effectiveness of human intervention in continuous process control [43]. Human Reliability Analysis (HRA) highlights how the *"human component"* affects the overall performance and reliability of heterogeneous systems [29].

Technology involves an extent of *risk* [50], regardless our knowledge or trust in it. Any time we use or rely on technologies we take risks. Understanding trust is very important in presence of uncertainties with respect to computer systems and, generally speaking, socio-technical systems. On the one hand technology supports human activities. On the other hand it is a source of harm. Engineering safety-critical systems involves risk analysis [34,55] as part of safety analysis in order to identify safety requirements, although assessing the benefits of technology exposes the limits of pure technical arguments [25]. Whatever is the risk associated with technology, social aspects constrain risk perception - *"Acceptable risk is a matter of judgement"* [10]. However, social and cultural aspects affect judgement [10]. For instance, MacKenzie analyzes how social connectivities affect global financial markets [37]. In particular, the study highlights how, even, electronic mediated trading relies on trust between traders communicating by computers [37]. This further points out contingencies between cooperation (competition) [44] and emergent trust (mistrust).

This paper analyzes trust in the context of Air Traffic Management (ATM). The future development of ATM, set by the ATM 2000+ Strategy [13], involves a structural revision of ATM processes, a new ATM concept and a systems approach for the ATM network. Despite the overall objectives [13], emerging lack of trust may undermine any improvement in the aviation domain (e.g., increased safety and performance). Ongoing research (see, Section 2) is debating and addressing the notion of trust: *What is trust? How to model trust?* Section 2 acknowledges that it is important to understand trust. But, it argues, too, that it is important to investigate trust dynamics. *Trust strategies* and *policies* should capture how socially constructed risk and knowledge (e.g., system reliability) interact each other. This paper stresses trust strategies (in terms of game theory) and trust policies for the investigation of interaction between *trust*, *risk* and *knowledge*. This paper is structured as follows. Section 2 reviews models of trust. Section 3 highlights the current developments in ATM. Section 4 introduces trust games and elaborates the motivations for trust strategies and policies in ATM.

[1] "Obviously, getting this process in place was difficult, for it involved major changes in the culture, especially when one pilot was junior. After all, when one person questions another's behavior, it implies a lack of trust; and when two people are supposed to work together, especially when one is superior to the other, trust is essential. It took a while before the aviation community learned to take the questioning as a mark of respect, rather than a lack of trust, and for senior pilots to insist that junior ones question all of their actions. The result has been increased safety", p. 145, [46].