Toward a Normative Approach for Resilient Multiagent Systems: A Summary

JAAMAS Track

Geeta Mahala
University of Wollongong
Wollongong, Australia
gm168@uowmail.edu.au

Ozgur Kafali
University of Kent
Canterbury, United Kingdom
ozgurkafali@gmail.com

Hoa Khanh Dam
University of Wollongong
Wollongong, Australia
hoa@uow.edu.au

Aditya Ghose
University of Wollongong
Wollongong, Australia
aditya.ghose@gmail.com

Munindar P. Singh
North Carolina State University
Raleigh, USA
mpsingh@ncsu.edu

ABSTRACT

We model a multiagent system (MAS) in socio-technical terms, combining a social layer consisting of norms with a technical layer consisting of actions that the agents execute. We express stakeholder needs to ensure that a MAS demonstrates resilience, allowing it to recover effectively from failures within a brief timeframe. This extended abstract presents a framework that computes probabilistic and temporal guarantees on whether the underlying requirements are met or, if failed, recovered. An important contribution of the framework is that it shows how the social and technical layers can be modeled jointly to enable the construction of resilient systems of autonomous agents. This paper facilitates specification refinement through methodological guidelines, emphasizing joint modeling of social and technical layers. We demonstrate our framework using a manufacturing scenario with competing public, industrial, and environmental requirements. This is an extended abstract of our JAAMAS paper available online [11].

KEYWORDS

Norms; Resilience; Formal verification

ACM Reference Format:
Geeta Mahala, Ozgur Kafali, Hoa Khanh Dam, Aditya Ghose, and Munindar P. Singh. 2024. Toward a Normative Approach for Resilient Multiagent Systems: A Summary: JAAMAS Track. In *Proc. of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024), Auckland, New Zealand, May 6 – 10, 2024*, IFAAMAS, 3 pages.

1 INTRODUCTION

Models of social interaction are central to artificial intelligence (AI), especially in the field of MAS [1, 5, 12, 14]. Socio-technical system (STS) serves as governance mechanisms in MAS [3, 9], where autonomous agents represent stakeholders' needs [2, 4, 6]. We adopt a conception of an STS [7, 8], where agents interact within the technical architecture. This STS is represented as a MAS and governed by norms [15] that regulate interactions, guiding agents toward fulfilling stakeholders' needs.

Motivating example. Consider a personal protective equipment (PPE) manufacturing scenario where different stakeholders have potentially conflicting functional and sustainability requirements. A textiles *company* is committed to meeting *hospital* demand by producing PPE at a reasonable price. At the same time, the *company* is prohibited by the *regulator* from polluting the environment. If violated, this prohibition may result in fines for the company and possibly a revocation of its permit to operate. Each stakeholder has a set of alternative actions; in general, these actions, affect the satisfaction and violation of the applicable norms differently. For example, the *company* can produce PPE in a sustainable manner, which reduces pollution but increases cost. Or, the *company* can produce cheap PPE, which reduces cost but increases pollution and the risk of being fined by the regulator. The way the company produces PPE and deals with the waste as a result of production also determines how resilient the overall STS is, e.g., how fast pollution can be reduced if it goes above a certain level.

This paper introduces the concept of resilient socio-technical systems, which aim to meet stakeholder requirements and recover from failures in meeting those requirements. Resilience becomes a crucial aspect of an STS's trustworthiness, particularly concerning its ability, as highlighted by [13]. We define the resilience of an STS based on the following key criteria: the ability to recover (1) from an undesirable state, (2) within a specified deadline or number of steps, and (3) with a probability exceeding a defined threshold. To operationalize these criteria, we extend STS specifications to incorporate time and quantities. Our extended paper [11] proposes a formalization and algorithm to translate STS specifications into PRISM (Probabilistic Symbolic Model Checker) [10] compatible models for model-checking. It introduces probabilistic model-checking for assessing resilience requirement probabilities, enabling trade-off evaluations between technical objectives and social regulations.

2 THE PROPOSED FRAMEWORK

Figure 1 illustrates the main components of the framework.

Stakeholders Stakeholders define the requirements that the STS must satisfy. If it appears that meeting a specific set

Proc. of the 23rd International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2024), N. Alechina, V. Dignum, M. Dastani, J.S. Sichman (eds.), May 6 – 10, 2024, Auckland, New Zealand.

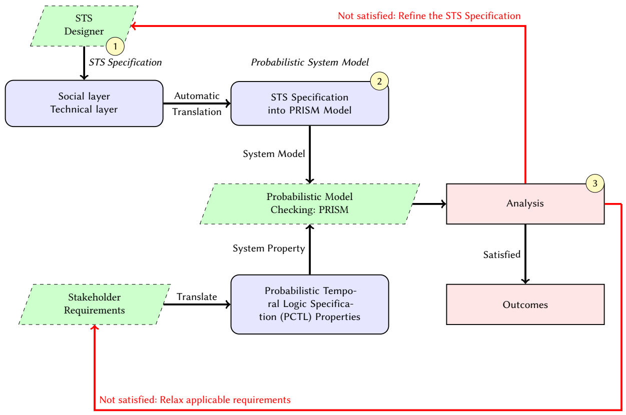

Figure 1: The Proposed Framework.

of requirements is not possible, stakeholders may need to reconsider and modify their requirements to create a feasible design for the STS. The rationale for this adjustment, such as having too few steps, is difficult to predict in advance. This aspect is addressed through an iterative process outlined in the full paper [11] under methodological guidelines.

We introduce *resilience requirements* to quantify an STS's ability to recover and adapt to adverse conditions. The equation below outlines a generic resilience requirement in the PCTL (Probabilistic Computation Tree Logic) syntax. It specifies that if the system enters an undesirable state ($UState$), it must return to a desirable state ($DState$) within a defined number of steps (stepConst) with a given probability (probConst), based on a specified operator ($\langle ineq \rangle$).

$$\langle UState\rangle \rightarrow P_{\langle ineq\rangle \text{probConst}}[F^{\langle ineq\rangle \text{stepConst}}\langle DState\rangle]$$

STS Designer The designer outlines an STS with two layers. The *social layer* defines the agents and the associated norms, while the *technical layer* specifies the actions undertaken by these agents. The social layer consists of a set of norms that govern the interactions among the agents. Note that norms [16] in our model are directed from one party to another. In our example, each PPE manufacturer would commit to the *Regulator* to meet certain pollution standards. That is, the norms are pairwise. Actions in the technical tier allow or restrict specific agent actions as they represent hard constraints. Actions describe relevant facts about the operating environment, e.g., what happens when an action is executed.

Translation The framework utilizes the STS specification as an input to create a PRISM model, which is a probabilistic state transition model. In this context, a probabilistic state transition model assigns a specific *transition probability* to each transition. What is generated is a slight variation known as *augmented probabilistic state transition models.* These augmented models consider both the probability of selecting an action and the probability of executing that action. The algorithm for translating an STS specification into a PRISM model is explained in the full paper [11].

Analysis The framework assesses the STS by comparing it to specified requirements and computing the likelihood of the STS violating or meeting each requirement. If improvement is needed, the STS designer enhance the STS specification. If modification is deemed unfeasible, stakeholders explore relaxed requirements through state variables or parameters.

2.1 Social Layer

In the spirit of Kafalı et al. [8] and Singh [15], a *norm* is defined as $\langle n, \text{SBJ}, \text{OBJ}, \text{ant}, \text{con}\rangle$. Here, $n \in \{c, p\}$ represents commitment or prohibition, SBJ and OBJ are subject and object from a set of agents $\mathbb{AG}$, and ant/con are conditions denoting antecedent/consequent. In Listing 1, a commitment c (Company, Hospital, true, PPE $\geq$ 100) means the Company commits to the Hospital to consistently produce more than 100 units of PPE. On the other hand, a prohibition (p) such as p (Company, Regulator, true, pollution $\geq$ 60) implies the Regulator prohibits the Company from exceeding 60 ppm pollution.

2.2 Technical Layer

The technical layer consists of a finite set of operational actions, denoted as $A = a_1, a_2, \ldots, a_k$, that agents can execute. Each action, represented as a (condition, DeleteList, AddList), signifies a state transition. Upon satisfying the condition, the action is executed, updating attributes by removing specified values in DeleteList and adding new values from AddList. In Listing 1, executing action $a11$ results in a new state where the variable PPE ranges from 50 to 100, and the variable *pollution* is the product of PPE and a step-size in [0.2, 0.4].

Listing 1: STS specification for PPE manufacturing.

```
p(Company, Regulator, true, pollution≥60)
c(Company, Hospital, true, PPE≥100)

a11: a(true, {PPE, pollution}, {PPE+=[50, 100],
        pollution+=PPE*[0.2, 0.4]})
a12: a(true, {PPE, pollution}, {PPE+=[80, 120],
        pollution+=PPE*[0.5, 0.7]})
```

3 CONCLUSIONS

We present a novel probabilistic framework for designing and verifying STSs that incorporates social norms, technical actions, and probabilistic temporal stakeholder requirements. Our key contribution lies in integrating socio-technical resilience and probabilistic model checking into a comprehensive methodology for specifying and verifying STSs. We have integrated norms into the PRISM model by assessing the likelihood of an action's execution, considering how its outcome aligns with the norms in the STS.

ACKNOWLEDGMENT

We dedicate this article to the memory of Professor Aditya Ghose, who passed away unexpectedly in February 2023.

This work was initiated with support from a grant from the University of Wollongong and NC State University through a collaboration network called the University Global Partnership Network.

MPS additionally thanks the US National Science Foundation (grant IIS-1908374) for partial support.

REFERENCES

[1] Cristiano Castelfranchi. 1998. Modelling Social Action for AI Agents. *Artificial Intelligence* 103, 1–2 (Aug. 1998), 157–182. https://doi.org/10.1016/S0004-3702(98)00056-3

[2] Amit K. Chopra, Fabiano Dalpiaz, F. Başak Aydemir, Paolo Giorgini, John Mylopoulos, and Munindar P. Singh. 2014. Protos: Foundations for Engineering Innovative Sociotechnical Systems. In *Proceedings of the 22nd IEEE International Requirements Engineering Conference (RE)*. IEEE Computer Society, Karlskrona, Sweden, 53–62. https://doi.org/10.1109/RE.2014.6912247

[3] Amit K. Chopra and Munindar P. Singh. 2018. Sociotechnical Systems and Ethics in the Large. In *Proceedings of the AAAI/ACM Conference on Artificial Intelligence, Ethics, and Society (AIES)*. ACM, New Orleans, 48–53. https://doi.org/10.1145/3278721.3278740

[4] Fabiano Dalpiaz, Paolo Giorgini, and John Mylopoulos. 2013. Adaptive sociotechnical systems: A requirements-based approach. *Requirements Engineering* 18, 1 (2013), 1–24. https://doi.org/10.1007/S00766-011-0132-1

[5] Mark d'Inverno, Michael Luck, Pablo Noriega, Juan A. Rodríguez-Aguilar, and Carles Sierra. 2012. Communicating Open Systems. *Artificial Intelligence* 186 (July 2012), 38–94. https://doi.org/10.1016/j.artint.2012.03.004

[6] Andrew J. I. Jones, Alexander Artikis, and Jeremy Pitt. 2013. The Design of Intelligent Socio-Technical Systems. *Artificial Intelligence Review* 39, 1 (2013), 5–20. https://doi.org/10.1007/s10462-012-9387-2

[7] Özgür Kafalı, Nirav Ajmeri, and Munindar P. Singh. 2016. Revani: Revising and Verifying Normative Specifications for Privacy. *IEEE Intelligent Systems (IS)* 31, 5 (Sept. 2016), 8–15. https://doi.org/10.1109/MIS.2016.89

[8] Özgür Kafalı, Nirav Ajmeri, and Munindar P. Singh. 2020. Specification of Sociotechnical Systems via Patterns of Regulation and Control. *ACM Transactions on Software Engineering and Methodology (TOSEM)* 29, 1 (Feb. 2020), 7:1–7:50. https://doi.org/10.1145/3365664

[9] Timotheus Kampik, Adnane Mansour, Olivier Boissier, Sabrina Kirrane, Julian Padget, Terry R. Payne, Munindar P. Singh, Valentina Tamma, and Antoine Zimmerman. 2022. Governance of Autonomous Agents on the Web: Challenges and Opportunities. *ACM Transactions on Internet Technology (TOIT)* 22, 4, Article 104 (Nov. 2022), 31 pages. https://doi.org/10.1145/3507910

[10] Marta Kwiatkowska, Gethin Norman, and David Parker. 2002. PRISM: Probabilistic Symbolic Model Checker. In *Proceedings of the International Conference on Modelling Techniques and Tools for Computer Performance Evaluation*. Springer, London, 200–204. https://doi.org/10.1007/3-540-46029-2_13

[11] Geeta Mahala, Özgür Kafalı, Hoa Dam, Aditya Ghose, and Munindar P Singh. 2023. A normative approach for resilient multiagent systems. *Autonomous Agents and Multi-Agent Systems* 37, 2 (2023), 46.

[12] Timothy J. Norman and Chris Reed. 2010. A logic of delegation. *Artificial Intelligence* 174, 1 (Jan. 2010), 51–71. https://doi.org/10.1016/j.artint.2009.10.001

[13] Amika M. Singh and Munindar P. Singh. 2023. Wasabi: A Conceptual Model for Trustworthy Artificial Intelligence. *IEEE Computer* 56, 2 (Feb. 2023), 20–28. https://doi.org/10.1109/MC.2022.3212022

[14] Munindar P. Singh. 1991. Group Ability and Structure. In *Decentralized Artificial Intelligence, Volume 2*, Yves Demazeau and Jean-Pierre Müller (Eds.). Elsevier/North-Holland, Amsterdam, 127–145. Revised proceedings of the *2nd European Workshop on Modeling Autonomous Agents in a Multi Agent World (MAAMAW)*, St. Quentin en Yvelines, France, August 1990.

[15] Munindar P. Singh. 2013. Norms as a Basis for Governing Sociotechnical Systems. *ACM Transactions on Intelligent Systems and Technology (TIST)* 5, 1 (Dec. 2013), 21:1–21:23. https://doi.org/10.1145/2542182.2542203

[16] Munindar P Singh. 2014. Norms as a basis for governing sociotechnical systems. *ACM Transactions on Intelligent Systems and Technology (TIST)* 5, 1 (2014), 1–23. https://doi.org/10.1145/2542182.2542203