BRIEFING NOTES

#BN-51-Emerging technology and military application-Feb2021

PUBLIC POLICY CHALLENGES, REGULATIONS, OVERSIGHT, TECHNICAL, AND ETHICAL CONSIDERATIONS FOR AUTONOMOUS SYSTEMS (AS): PART II

Authors: Neshat Elhami Fard [1], Kianoush Haratiannejadi [1], Rastko R. Selmic [2], Khashayar Khorasani [3]

[1] Graduate students, Department of Electrical and Computer Engineering, Concordia University, Montreal, Canada

[2] Professor, Department of Electrical and Computer Engineering, Concordia University, Montreal, Canada; rastko.selmic@concordia.ca

[3] Professor, Department of Electrical and Computer Engineering, Concordia University, Montreal, Canada; kash@ece.concordia.ca

SUMMARY

- Autonomous systems (AS) are systems with a variety of sensors to understand environmental information so that they can distinguish, evaluate, and make decisions based on them [1]. In addition to an autonomous single-agent system, the AS can be designed in the form of multi-agents to identify high-risk, hazardous, or inaccessible areas [2], [3].
- Autonomy in systems is directly related to security of such systems. If an AS is compromised, virtual and physical problems occur that lead to loss of data, communication interruptions, and damage or loss of the system [4]. To ensure a proper operation, robot operating system (ROS) security features are optimized utilizing encrypted communications and semantic rules [5].
- Decision-making in all types of AS is another significant issue. An always present dilemma -- ``which decision should be made by a software system, and which one should be made by a human." Software developers for AS have a responsibility to answer the above question [6].

CONTEXT

Levels of Autonomy:

Human-Robot Interaction (HRI) defines the types of robot interactions, with each category requiring a different level of autonomy. Developers should evaluate levels of autonomy suitable for their robot according to the framework guidelines, and investigate the effects of autonomy on HRI [7]:

- According to SAE International's new standard J3016, there are six levels of driving automation for an autonomous vehicle (AV), namely level 0: no automation, level 1: driver assistance, level 2: partial automation, level 3: conditional automation, level 4: high automation, and level 5: full automation. At levels 0, 1, and 2, a human driver monitors the driving environment, while at levels 3, 4, and 5, an automated driving system performs surveillance [8].
- Aeronyde defines drones with seven levels of autonomy. Level 0: no automation, level 1: pilot assistance, level 2: partial automation, level 3: conditional automation, level 4: high automation, level 5: adaptive autonomy, and level 6: augmented autonomy [9].
- The Air Force Research Laboratory has outlined levels of autonomy for different systems. In this regard, Los Alamos National Lab has described six autonomy levels for MAP (mobility, acquisition, and protection) survival space. Level ``-": motion only occurs under application of an external force, level 0: no motion abilities, level 1: moves independently in one dimension, level 2: moves deliberately in two

dimensions, level 3: moves independently in tree dimensions, level 4: capable of dual-mode motion with tools, vehicles, or application of specific design elements, and level 5: human [10].

- Charles Stark Draper Laboratory, which has been designing and developing robotic systems for the military for many years, classified the autonomy levels of Draper 3D Intelligence Space. According to this classification, level 1: no mobility control-robotic process automation (RPA) only, level 2: operator-assisted, level 3: get to the waypoint and do one feature-based command, and finally level 4: integrate multiple actions [10], [11].
- Based on the first general classification of autonomous control level (ACL), the eleven levels (0-10) of Clough's autonomy control level framework are as follows, respectively: a remotely piloted vehicle executes a pre-planned mission, a changeable mission, a robust response to real-time faults/events, and fault/event adaptive vehicle. The rest are real-time multi-vehicle coordination, real-time multi-vehicle cooperation, battle space knowledge, battle space cognizance, battle space swarm cognizance, and fully autonomous [10].
- The United States Army Science Board has classified the level of autonomy differently. Level 0: manual remote control, such as a remote-controlled toy; level 1: simple automation; level 2: automated tasks and functions, such as a hunter; level 3: scripted mission, such as a shadow or predator unmanned aerial vehicle (UAV); level 4: semi-automated missions with simple decision-making, such as a cruise missile; level 5: complex missions-specific reasoning; level 6: dynamically mission adaptable; level 7: synergistic multi-mission logic; level 8: human-like autonomy in a mixed team; level 9: autonomous teams with unmanned leader or mission manager; level 10: autonomous conglomerate. In the aforementioned list, the complexity, capability, flexibility, and adaptability enhance with increasing level [12].

Ensuring Trust and Reliability in Autonomous Systems: Trust is one of the most significant issues in designing semi-AS and AS. An AS that is not trustworthy cannot be used. Hence, to have trust, an AS must consider four high-quality factors. These vital services that affect the trustworthiness of AS are (i) robust for any health concerns, (ii) safe for any stuff in their surrounding environments, (iii) secure for any threats from cyberspaces, and (iv) reliable for human-machine interaction.

Factors Affecting Trust: The factors that affect AS's trust have generally been classified into characteristics of the system, characteristics of the operator, and the environment's characteristics.

System Properties: The essential correlations of AS use have been system reliability and the effects of system faults. Reliability refers to AS that has some error rate, for example, misclassifying targets. Faults are harsher and have broader impacts, causing the AS to act carelessly without paying attention to the circumstances [13].

- **System Reliability:** Studies in system reliability show that one reason the user loses his trust is declining system reliability. The system reliability and the user's trust in that AS can be measured over time [14].
- **System Faults:** System fault is a subsection of system reliability. Different aspects of faults affect the connection between trust and AS. In [14], the author pointed out that trust in the AS decreases in the presence of frequent system faults. However, the trust can be recovered slowly, even as faults resumed, but it has to be controllable. The size of system faults and its repetition have differential impacts on trust. Faults of differing sizes decreased trust more than repeated faults.
- **System Predictability:** Predictability of AS causes users to gain trust in the system. Even if a system fault influences and decreases AS's trust, the predictability feature improves the trust. Different studies have shown that when people have prior awareness of the possible faults or malfunctions, they do not necessarily lose their trust in the AS [15].
- **System Intelligibility and Transparency:** The most important feature of the AS system is the ability to justify its decision and its logic behind that. This feature makes the AS more liable to be trusted since its users understand why specific decision has been made [16].

Properties of the Operator: A user may trust AS in general; however, it does not mean that he/she trust all of their applications [17]. For instance, the user might trust a self-driving car, but he should not necessary trust to an autonomous weapon system (AWS). Moreover, the user's beliefs are one of the operator's crucial characteristics that have a significant effect on the AS's trust. It is essential to consider the competence of the user that is judging the AS. For instance, a civilian user's opinion might not be as important as a military commander when it comes to judging an AWS in a warzone.

Environmental Factors: In terms of environmental factors that affect AS trust, the risk that originates from the environment appears most important. Research in AS trust implies that one of the factors that affect the AS's confidence is the risk present in the decision-making process related to the AS's environment. Another study shows that once trust has been eliminated or decreased, it requires to reach higher confidence for the user to use that AS in

high-risk environments. However, knowing the AS failure risk in that environment in advance may decrease the user's distrust in the AS [13], [18].

CONSIDERATIONS

- Examples of dual-use technologies and threats they may pose to the public safety [2].

Opportunities and Risks of Autonomous Systems for Royal Netherlands Army (RNLA)

The opportunities are:

- Generating more reliable and quicker situational consciousness and perception.
- Extending the ability, persistence, and stability of operations.
- Diminishing the physical and cognitive loads of soldiers.
- Allowing and enabling the simultaneous execution of tasks for effective and useful actions.
- Enhancing the speed of the OODA (Observe, Orient, Decide, Act) loop.
- Defending and protecting the force.

The risks are:

- Communication signals applied by AS are vulnerable and assailable to cyberattacks, namely hacking, blocking, and disrupting system performance.
- How data is interpreted, and practical information is provided, is extremely tough for AS to simplify trust in decision-making.
- How AS work, and how to solve their upcoming problems, is incomprehensible for defense communities.
- Lack of development of various machines and moving towards autonomy due to insufficient trust in AS.
- Operators' overconfidence on AS.
- Creating possible new tensions between traditional soldiers and new technical experts and data scientists.
- Changing of training requirements, education, careers, and the type of work that soldiers engage in, as well as, exchanging leadership positions due to generating AS.
- The integration of AS and the possibility of replacing machines in some individuals or units will have an impact on the training of the armed forces.
- In some AS, such as the killer robot, a negative public perception leads to a lack of consideration for the benefits of automation and autonomy and thus considering the human control.
- Authoritarian and rebellious governments that care less about moral considerations may reinforce the AS and use them in dangerous inhumane ways.

- The problem and challenge of adapting international and national laws to the creation and development of autonomous technologies.
- Since those who use AS are able to distract and deny responsibility for attacks, assigning individuals to use AS is a challenging task.
- Opponents and adversaries may have many moral and legal restrictions on the reproduction and use of AS in all areas.
- The creation of AS may lead to an arms race so that powerful countries can show and use their potential.

Opportunities and Risks of Autonomous Systems in Logistics

The opportunities are:

- Creating new jobs for the elderly, people with disabilities, and the underprivileged by integrating them with AS, especially for jobs like autonomous truck driver.
- Reducing shipping operation costs up to 40% by eliminating drivers and using autonomous vehicles.
- Save fuel by creating a network, including a leader and several followers. The leader determines the speed and direction of the rest of autonomous vehicles, thereby preventing additional fuel consumption.

The risks are:

- Tendency to make mistakes, errors, and possible situational misjudgments because of their computer nature, and consequently leaning to accident.
- Increasing job loss, and consequently high unemployment rate.

Opportunities and Risks of Autonomous Systems in Healthcare

The opportunities are:

- Opening up new commercial opportunities for the insurance and healthcare industries.
- Reduction of emergencies leading to death for high-risk patients due to active monitoring.

The risk is:

- Rising privacy and security concerns due to weak security structure.

NEXT STEPS

In future, the appropriate level of human involvement in AS will be investigated.

REFERENCES

[1] H. Xinhan and W. Min, "Multi-sensor data fusion structures in autonomous systems: a review," in Proceedings of the 2003 IEEE International Symposium on Intelligent Control. IEEE, 2003, pp. 817–821.
[2] W. Fink, J. Dohm, and M. A. Tarbell, "Multi-agent autonomous system," Jan. 24 2006, US Patent 6,990,406.
[3] W. Fink, J. Dohm, and M. A. Tarbell, "Multi-agent autonomous system and method," June 22 2010, US Patent 7,742,845.
[4] F. J. R. Lera, C. F. Llamas, A. M. Guerrero, and V. M. Olivera, "Cybersecurity of robotics and autonomous systems: Privacy and safety," Robotics-legal, ethical and socioeconomic impacts, 2017.
[5] J. Balsa-Comeron, A. M. Guerrero-Higueras, F. J. Rodriguez-Lera, C. Fernandez-Llamas, and V. Matellan-Olivera, "Cybersecurity in autonomous systems: hardening ROS using encrypted communications and semantic rules," in Iberian Robotics Conference. Springer, 2017, pp. 67–78.
[6] T. Linz, "Testing autonomous systems," in The Future of Software Quality Assurance. Springer, Cham, 2020, pp. 61–75.
[7] J. M. Beer, A. D. Fisk, and W. A. Rogers, "Toward a framework for levels of robot autonomy in human-robot interaction," Journal of humanrobot interaction, vol. 3, no. 2, p. 74, 2014.
[8] S. International, "Automated driving levels of driving automation are defined in new SAE international standard j3016," 2014.
[9] "Self-flying drones: Who will be in the pilot's seat?" https://aeronyde.com/2018/08/28/2018-8-28-self-flying-drones-whowill-be-in-the-pilots-seat-1/, Aug. 2018.
[10] B. T. Clough, "Metrics, schmetrics! how the heck do you determine a UAV's autonomy anyway," Air Force Research Lab Wright-Patterson AFB OH, Tech. Rep., 2002.
[11] M. E. Cleary, M. Abramson, M. B. Adams, and S. Kolitz, "Metrics for embedded collaborative intelligent systems," NIST SPECIAL PUBLICATION SP, pp. 295–301, 2001.
[12] P. Mulgaonka, J. Blair, R. Dodd, D. Martinez, and R.-D. J. Perna, "Ad hoc study on human robot interface issues," ARMY SCIENCE BOARD WASHINGTON DC, Tech. Rep., 2002.
[13] M. Lewis, K. Sycara, and P. Walker, "The role of trust in human-robot interaction," in Foundations of trusted autonomy. Springer, Cham, 2018, pp. 135–159.
[14] R. Haslam, A. Thatcher, and R. Haslam, "Editorial–tribute to Neville Moray (1935–2017)–ergonomics and global issues," 2018.
[15] P. Jandric, "Post-truth and critical pedagogy of trust," in Post-Truth, Fake News. Springer, 2018, pp. 101–111.
[16] B. French, A. Duenser, and A. Heathcote, "Trust in automation," 2018.
[17] J. D. Lee, "Human factors: The journal of the human factors and ergonomics society, Jan. 1, 2002."

[18] V. Riley, “Automation theory and applications, chapter operator reliance on automation: theory and data,” 1996.
[19] B. Torossian, F. Bekkers, T. Sweijs, M. Roelen, A. Hristov, and S. Atalla, The Military Applicability of Robotic and Autonomous Systems. Hague Centre for Strategic Studies, 2020.
[20] M. Grolms, “Autonomous vehicles in logistics part 1: Opportunities and risks,” https://www.allthingssupplychain.com/autonomous-vehiclesin-logistics-part-1-opportunities-and-risks/, Jun 2020.
[21] O. David-West, “What do new autonomous technologies mean for global business?” https://globalnetwork.io/perspectives/2016/09/what-do-newautonomous-technologies-mean-global-business, Sep 2016.