Human Factors and Systems Interaction, Vol. 52, 2022, 41–47
https://doi.org/10.54941/ahfe1002137

Envisioning 21st Century Mixed-Initiative Operations for Energy Systems

Roger Lew[1], Ronald L. Boring[2], and Thomas A. Ulrich[2]

[1]University of Idaho, Moscow, ID, USA
[2]Idaho National Laboratory, Idaho Falls, ID, USA

ABSTRACT

Human technological systems are growing increasing complex and sophisticated. New energy opportunities are emerging with the proliferation of utility and consumer scale solar power that enables individuals to become both producers and consumers (prosumers) of power. Grid storage in the form of batteries, pumped hydro, and thermal are also becoming increasing capable and available for providing grid stability and resiliency. Artificial intelligence is quickly becoming cheaper to train and more capable at solving problems at scale to the point where it can be deployed and entrusted with critical infrastructure tasks. Taken together the electric grid is a complex system of systems that is partially evolved and partially designed, it is operated by human and machine agents. Determining how to maintain effective oversight of systems is a technical and philosophical challenge. Here we explore some qualitative classifications to guide those decisions.

Keywords: Nuclear power, Artificial intelligence, Mixed-initiative systems, Prosumers, Solar power, Grid storage, Smart gird

INTRODUCTION

In 1992 a computerized dispatch system for the London metropolitan area was developed and deployed for the London Ambulance Service (LAS). The LAS was responsible for providing service to 6.8 million people living in a 600 square mile area. On a typical day they would average between 2000 and 2500 calls to be handled by a fleet including 445 ambulances distributed across 70 stations. Prior to the computer-based system the process of dispatching ambulances was entirely manual. The new computer-based system was intended to reduce response time to within three minutes as required by recent legislation. The system included automatic vehicle locating system (AVLS) to automate resource identification and mobilization. The computerized LAS went live on October, 26, 1992 to service all 6.8 million residents. The deployment was a spectacular failure resulting in multiple units being sent to some locations while neglecting others. The automated system also dispatched units from far away when closer units were available. The problem was further exacerbated by a frantic public making multiple calls to ensure service. By the next-day the LAS reverted to a partially-manual system, and within 8-days the system completely stopped working and was shutdown. In the aftermath the system was attributed as causing as many as 46 deaths (Musick, 2006).

The LAS failed for numerous reasons: the hardware was under specified for the application, the development went to a low bidder who underestimated the work required and lacked the necessary skills to carry out the project. The system was also not properly commissioned and stress tested. Notably, the software contained a memory leak that ultimately crippled the system (Musick, 2006). We can discuss LAS as a warning to placing too much trust in machines and to point out the resilience of human decision makers. However, we can also note the success of transportation as a service companies like Uber, Lyft, and Waymo who have all demonstrated success as astounding scales.

Useful Artificial Intelligence

Fast-forward to 2022 and Uber algorithmically routes 93 million monthly users with 3.5 million drivers. In addition, Uber Eats generated $4.8 billion in revenue for 2022 implying Uber delivered hundreds of millions of meals. Jeff Clune, the former director of Uber's AI lab described "machine learning as the heart of almost every aspect of Uber." Artificial Intelligence is used to optimize ride sharing (where more than one set of passengers share a vehicle), to identify fake accounts, improve routing, and provide better prediction to volatility around historical events. Over the last decade the capabilities of artificial agents and our ability to deploy them at scale has greatly enhanced. This isn't to say they are perfect. Detractors of Uber point out that the overuse of AI for managing "gig workers" has lead to negative sentiments among drivers who can't help but feel they are being managed by machines who sometimes whimsically decide their livelihoods (Ma, Yuan, Ghafurian, Hanrahan, 2018).

Clune (2020) describes how tackling applied problems in computer science has evolved. Computer scientists use to decompose problems and design hand-coded solutions to problems. We problems proved too challenging we then hand-coded systems and used machine learning to optimize the systems. In some cases, we would hand-code the features, like edges, and gradients that the systems should learn. Then we found that the end results were better if the machine learning identified features on its own. Now we progressed to allowing machine learning to learn the system. Our understanding of how to train AI is also evolving. The traditional approach is to optimize for a single fitness function, however Clune illustrates with a trivial example of an agent minimizing distance to a target in a 2d maze why this is not satisfactory compared to an agent how is optimized to explore the space. An algorithm trained to minimize distance falls into local minima and fails to reach the target. The algorithm tasked with exploring the space reaches the target by chance. Clune also describes how with traditional machine learning we would optimize two variables (cost and time) by devising a single fitness function and allowing the algorithm to try to simultaneously optimize both. In contrast the two-dimensional reward landscape can be explored resulting in algorithm that has overall better performance. Meta-learning AI has found that goal shifting is key to success. By changing goals algorithms continue to adapt. These adaptations may not always lead to better outcomes with the current goal, but might prove useful for other goals or tasks. Likewise, changing the environment is also important to meta-learning. POET the Paired open-ended trailblazer algorithm trained a bipedal two-dimensional robot to walk. The presence of small blocks was instrumental to the robot walking in an upright fashion. The upright behavior did not evolve when the robot as on a flat surface. By changing the learning environments, the robot was eventually able to navigate heterogeneous terrain with steep slopes, pits, and other obstacles.

Key to the success of AI in the last decade is the pace of compute. Specifically, the cost to train AI is improving at 50x the pace of Moore's Law with training compute in Petaflops/day increasing at a pace of 10x a year (Wang, 2020). The future is AI applications becoming both ubiquitous and invisible. For instance, Adobe Photoshop is distinguished by top-tier AI that aides users in edge selection, intelligent masking and replacement, and photographic enhancements. But users of Photoshop are not trained as computer scientist – some might not even realize they are using AI.

Here the conjecture is that similar trends are inevitable with the eventual management, oversight and control of critical infrastructure. AI will offer increasing sophisticated and integrated capabilities to improve system reliability and resilience. Many might find this scenario unlikely, especially for risk adverse industries such as nuclear power and would even agree that operating paradigm for *existing* generation II reactors will not drastically change. However, the accelerating pace of technology is worth noting. Edison invented the lightbulb in 1880. EBR-1 was the first nuclear reactor to generate electricity in 1951. Our modern electric grid started as fragmented connecting power generators to commercial and residential consumers. Over time it has become more interconnected to the extent that it is now the largest most interconnected machine on Earth with over 9,000 generators and 1,000,000 megawatts of capacity (DOE, 2022). The most complex machine on Earth is continuing to evolve with real-time metering, supervisory control and data acquisition (SCADA) technologies, renewable energy sources and energy storage systems. The future of the gird is envisioned to be more intelligent to actively respond to overloads, accommodate energy from dynamic sources, while being reliable, resilient to failure, more efficient, and better for the environment (DOE, 2022). The electric grid is a prime example of a complex engineered system in that it is a system that is partially designed, and partially evolved. It is partially controlled by humans and partially by automation, and the configuration and dynamics of the system are dynamic.

Despite a slow pace, Nuclear Power is undergoing a global renaissance. Small modular reactors (SMR) and microreactors are in various design and commissioning phases. These are designed to be built in factories and installed onsite, providing a means to rapidly deploy nuclear power while controlling for uncertain capital expenditures and cost overruns. The OECD (2016) is projecting that by 2035 we could have 21 GWe of new nuclear electricity capacity installed globally with 3.5 GWe in the United States.

Simultaneously, renewables such as wind and solar are growing exponentially and battery electric vehicles are gaining traction in the energy sector. If vehicles transition to battery electric vehicles (BEV) our electricity consumption would roughly double. The energy grid as a whole is evolving as numerous point source generators come online and smarter grids enable better resource management and dynamic pricing. The result will be a distributed energy market where individuals and utilities both buy and sell resources in a fast-paced, brokered market. Or perhaps more accurately, autonomous agents will buy and sell resources on behalf of utilities, individuals, and intermediaries. The pertinent question then becomes how do we have human oversight of resources to maintain safe, secure, and reliable operation? Such a paradigm is a mixed initiative paradigm where both humans and autonomous agents monitor, supervise, and control systems.

MIXED INITIATIVE TEAMING

We live in an increasing interconnected world with exponentially increasing sensors. Coupled with commercial and consumer generation and storage there is an abundance of potential configurations to make for more efficient, less expensive, more reliable solutions. Making sense of the dynamic complexity and taking appropriate action becomes arguable the most important aspect of the problem.

Tasking AI with micro-second to second decisions to monitor distribution and generation and conduct automated switching is likely the only option. The alternative becomes having a human at a screen that when prompted complacently agrees and hoping the delay for the human approval wasn't too long.

Teamed decision making is well suited for tasks like forecasting, predictive maintenance, demand side management, and cybersecurity where the AI can remain vigilant, alert humans to potential issues, and conduct automated analyses and reports that can be interpreted by humans for decision making.

Product Risk Classifications

I think a reasonable approach is to examine assets as classes distinguished by risk. The least risky class comprises commodity consumer-oriented devices such as home photovoltaic, battery storage, and BEVs represented distributed nano-scale devices. The capital expenditures of any single device or installation are relatively small, and the potential consequences of a single installation failing are relatively small. Minimal regulatory oversight should be required for individual installations. These devices can charge and discharge autonomously with some high level rules for ensuring safety. For example, grid tie storage devices should be smart enough to not back feed power into the grid, and have expected features like current protection.

In composite these nano-scale devices could be operated in a synchronized manner to form a virtual power plant. Where the individual presence or absence of nano-scale devices have negligible impact to the grid as a whole, the synchronization of distributed assets could have meaningful impact on stabilizing the grid while reducing emissions. Such virtual power plants should have more regulatory oversight since there is potentially more impact (for better and worse) to the grid when resources are operated in a coordinated manner. The virtual power plants should have some human oversight and ability for human's taking manual control. The centralized coordination could also expose common cause failures or vulnerabilities (particularly cyber) that should be considered.

The second class comprises distributed micro-scale devices like nuclear micro-reactors and small modular reactors. These will have substantial automation compared to existing Generation II reactors. They could incorporate remote operations and monitoring at the fleet scale, with the ability to shut down systems locally. Disruptions would have costly impacts to an organization or municipality. Lastly, at the other end of the spectrum are high-value assets with the potential for low-probability high consequence events. These would include gigawatt-scale nuclear/solar/hydro plants that might also have flexible operations to support onsite data centers, hydrogen production, or cryptomining. These assets would be high-value targets and disruptions would have the potential for severe economic, environmental, and functional consequences at large geographic scales.

When we start thinking about human oversight, participation, and decision making, the first class is consumer-oriented. Consumers will be enabled to become prosumers (producers and consumers) sell excess or optimize energy usage and storage based on dynamic rates. The third class of high-value assets resembles how critical infrastructure is managed today. These high-value assets are conservative and slow to evolve through the adoption of automation and operational changes. They would still need to maintain high degrees of human vigilance compared to the other systems for regulatory adherence and maintaining cyber-physical security and reliability. The second class still has high regulatory requirements. However, it is a bit of a clean slate to conceptualize operations and monitoring from first principles with high levels of automation and mixed-initiative monitoring and control (AI/human teaming). In this paper we explore those possibilities. New SMR and microreactors incorporate passive safety and modern engineering modeling and analysis that wasn't available during the design and commisioning of Generation II reactors. The result is reactors that have significantly reduced risks of catastrophic melt-down events like Fukishima. This dramatically expands the possibilities for how they can be monitored and controlled. When we ponder what modern nuclear control rooms should look like we envision multiple operators monitoring dozens of screens to maintain situational awareness and readiness to respond at a moments notice. However, this is unlikely and perhaps even undersired. Once reactors, in particular microreactors, have the demonstrated capability of operating hands-free with minimal oversight it becomes misguided to install humans to maintain constant vigilance (e.g. Level 4 to 5 self-driving).

The key performance indicator should be system performance not situational awareness. Having "operators" permanently installed in a control room when no action is required 99.9% of the time becomes a superficial level of vigilance. Take system administration as a corollary. System administrator's primary responsibility is to maintain the availability of infrastructure, but their primary tasking is not to sit idly by and actively monitor. Human Factors Engineering involves understanding the need for comprehensive integration of human capabilities (cognitive, physical, sensory, and team dynamics) into a system design, beginning with conceptualization and continuing through system disposal. The primary concern for human factors engineering is the need to effectively integrate human capabilities with system interfaces to achieve optimal total system performance (use, operation, maintenance, support, and sustainment). Human factors engineering utilizes comprehensive task analyses to help define system functions and then allocates those functions to meet system requirements. The goal of HSI is to optimize total system performance, accommodating the characteristics of the user population that will operate, maintain, and support the system, and minimize life-cycle costs (Folds et al. 2008). HSI experts work within the Systems Engineering (SE) process to ensure that all human considerations are integrated throughout system design, development, fielding, sustainment, and retirement. The attention to human systems integration in system development programs drove hundreds of human-centered design improvements. Efforts were concentrated to maximize total system performance through improvements in human workload, ease of maintenance, and personnel safety which resulted in a cost avoidance of billions of dollars and prevention of hundreds of fatalities and disabling injuries for the system (Booher and Minninger, 2003).

CONCLUSION

Here we examine current trends in energy and artificial intelligence from the perspective of human oversight. The energy grid is a complex system of systems with increasing capable technologies for enabling a sustainable future. Challenges exist to leverage the vast potential of configurations and available data to form choices that optimize personal, societal, and environmental criteria. Artificial intelligence and automation will play a critical role in future systems and technology. Work is needed to determine how to pair human assets with machine intelligence.

REFERENCES

Clune, J. (2020). AI-GAs: AI-generating algorithms, an alternate paradigm for producing general artificial intelligence). arXiv:1905.10985v2 [cs.AI]

Dean, B. (2022). Uber Statistics 2022: How Many People Ride with Uber? BackLinko. https://backlinko.com/uber-users

DOE (2022). The smart grid an introduction. DE-AC26-04NT41817, Subtask 560.01.04.

Ma, N. F., Yuan, C. W., Ghafurian, M, Hanrahan, B. V. (2018). Using Stakeholder Theory to Examine Drivers' Stake in Uber. CHI 2018. doi:10.1145/3173574.3173657

Musick, E. (2006). The 1992 London Ambulance Service Computer Aided Dispatch System Failure. ErichMusick.com https://erichmusick.com/pdf/writings/technology/1992-london-ambulance-cad-failure.pdf

Novatio, Uber and Lyft Are Taking Artificial Intelligence Along for the Ride. Accessed February 10, 2022. https://www.novatiosolutions.com/uber-lyft-taking-artificial-intelligence-along-ride/

Wang, J. (2020). The Cost of AI Training is Improving at 50x the Speed of Moore's Law: Why It's Still Early Days for AI. Ark Invest. https://ark-invest.com/articles/analyst-research/ai-training/

Vorwerck. M. (2019). Advancing AI: A Conversation with Jeff Clune, Senior Research Manager at Uber. Uber Engineering. https://eng.uber.com/jeff-clune-interview/