Research Statement

Yinshuang Xiao

In the era of ever-increasing social and technological interconnectedness, complex socio-technical systems (STS) have permeated various facets of modern life. These systems encompass a broad range of applications, from large-scale service systems (e.g., smart power grids, shared mobility, and electric vehicle (EV) charging infrastructures) to socio-economic systems (e.g., customer-product market networks) and human-AI collaboration systems (e.g., smart manufacturing). However, **traditional methods face significant challenges when engineering and designing STSs for increasing and diverse human needs, particularly in three critical aspects**: 1) **Modeling:** The complexity of STSs, characterized by their scale, different stakeholders, intricate interdependencies, and temporal dynamics, makes effective modeling a difficult task. 2) **Analysis:** Existing methods struggle to efficiently address the forward problem, i.e., STS analysis, particularly in quantifying the intricate interdependencies between entities (e.g., neighboring charging stations sharing usage demand) and the uncertainties introduced by human factors and artificial intelligence (AI). 3) **Design:** There is a lack of effective approaches to tackle the inverse problem—designing STSs with desired features and performance while accounting for the interdependencies between system components. To address these challenges, there must be a paradigm shift in methodology and tools. To this end, my **long-term vision** is to establish the theoretical foundation for STS engineering and design to tackle the challenges associated with the aforementioned three aspects. I aim to build the foundation by synergistically integrating knowledge from disciplines in network science, machine learning, optimization, and statistics.

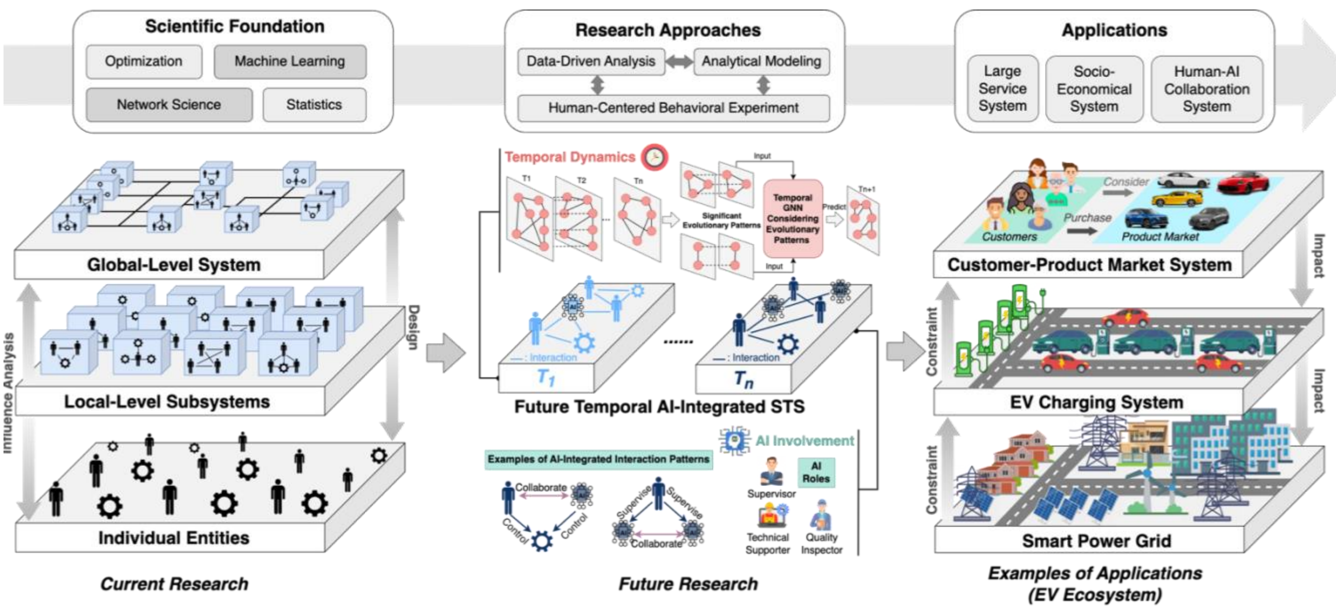

Figure 1: Research Overview

Research Experience and Contributions

As shown in Figure 1, the **central hypothesis** of my research is that the statistically significant local connections of individual entities, embedding collective behaviors of entities, are crucial function units of STSs, influencing global-level system performance. A paradigmatic example is the shared mobility network, such as bike-sharing systems. Compared to individual stations that only record users' rental and return behaviors, local service systems formed by several stations (e.g., a triangle of three stations) capture users' travel behaviors within the local network (e.g., a triangle transit). A better understanding of such a subsystem can help better probe into users' mobility patterns and local resource distribution, which are critical for the optimal decision-making of system engineers and operators. Despite the significance of local network structures and behaviors, existing work has primarily focused on investigating individual entities and the impact of their behaviors on system-level performance. Therefore, a fundamental knowledge gap exists in understanding and utilizing meaningful local network structures for STS engineering and design. My research has contributed to the current literature by creating a novel local-level network-based framework for STS engineering and design, combining advanced network theories and models, including network motif theory, exponential random graph model (ERGM), and graph neural network (GNN) models. Based on two real-world applications, i.e., customer-product market systems and shared mobility systems, I have demonstrated that the proposed framework can solve several critical challenges, such as modeling large-scale STSs with diverse interdependencies, identifying behavior patterns of system entities, and designing STSs with desired performance incorporating interdependencies among entities. In particular, this framework contains two phases. In Phase 1, I contributed to identifying entity behavior patterns by mining statistically significant local network motifs (i.e., recurring network patterns) and analyzing their impacts on system structure and performance. In Phase 2, I developed a new method combining the local information obtained from Phase 1, optimization design, and network predictive models (e.g., ERGM and GNN) to solve the inverse problem in STS engineering. Overall, **the unique contribution of the proposed framework** is that **it filled the fundamental knowledge gap in understanding meaningful local connections of individual entities** and, **for the first time, incorporated this local-level structural information into the design of STSs, leading to improved system performance**.

1. Phase 1: Significant Local Network Motifs Identification and Impact Analysis

The research outcome of Phase 1 is an analytical framework to address the challenge of extracting local network motifs embedded within the global network. In the case study on a bike-sharing system (Chicago's Divvy Bike), I developed a GNN-based model to predict whether a pair of stations would have sufficient travel demand. This model was found to have **a ~10% improvement** over traditional neural network models because of the incorporation of local network information in machine learning [1]. My first-authored paper in this line of work won the **2023 Journal of Mechanical Design Editors' Choice Honorable Mention**[1]. In another case study on customer-product market systems, I first defined a co-consideration network representation where nodes represent product models and links denote the co-consideration relations between two products (i.e., two products are co-considered at least once). This network model has been proven to be an effective representation of product competition, from which essential competition patterns were successfully identified. For example, I used ERGM and discovered **the pattern where products from three brands (i.e., inter-brand competition) form a closed triangular motif, positively influencing products' competitiveness.** This enables enterprises as market players to measure their competitive standing by quantifying the frequencies of their products engaged in this pattern [2].

2. Phase 2: Using The Identified Local-Level System Information to Guide STS Engineering and Design

This phase, for the first time, solves the inverse problem in STS by investigating how local-level system information from Phase 1 can guide STS design. In the bike-sharing case study, I discovered that the imbalanced number of rental and return bikes within a local service system (e.g., a local system consists of three stations) is sensitive to seasonal effects, e.g., serious imbalance issues occur more often in summertime, leading to user dissatisfaction. I further discovered a strong correlation (**~0.9** of the Person correlation coefficient) between this sensitivity and the differences in docking capacities among stations within those identified local systems. Based on these discoveries, I formulated an optimization problem to improve dock planning (i.e., how many docks shall be designed in a station) and seasonal robustness. This approach was successfully demonstrated in improving the robustness of Chicago's Divvy Bike against seasonal effects, reducing the largest dock capacity difference within local systems by **~5%** [3]. Moreover, in my recent publication about household vacuum cleaner market systems, I developed a network-based methodology to integrate local-level system information into the optimal design of vacuum cleaners' attributes. In this work, I first defined product competitiveness as a vector based on how often a vacuum cleaner model gets involved in unique competition patterns identified in Phase 1. This vector is essentially a function of a vacuum cleaner's attributes, such as suction power and weight, from which we can estimate the relationship between the competitiveness vector and the product's market share. Finally, I formulated an optimal design problem to find the new design attributes that maximize the product's market share while considering its competition relations. **The proposed method resulted in twice the market share increase of the targeted vacuum cleaner model compared to traditional optimal design approaches**, highlighting its potential for real-world application in product design [4].

Future Research

In the next five years, my **research objective** is to extend my current work by delving into STSs of greater complexity because of temporal effects and the embrace of AI technologies. This involves advancing our

[1] https://asmejmd.org/editors-choice-award/

understanding of the temporal dynamics of STSs and AI-integrated STSs (e.g., smart power grids) to inform effective engineering and design practices. In particular, I propose three research projects outlined below.

1. STS Engineering and Design Considering Temporal Dynamics

There is a consensus that STSs continuously evolve due to ever-changing human behaviors. Without a deep understanding of these dynamic changes, design strategies may become ineffective over time. For instance, in customer-product market systems, customer preferences shift, requiring design strategies that adapt to these changes to ensure future market success. Similarly, in power grid design, fluctuating energy demand is a key factor to consider for ensuring robustness and reliability. **Despite the importance of considering temporal changes in STSs, there is currently no efficient method to simulate these changes at both the system structure level** (e.g., the introduction of new EV models into the market) **and the individual attribute level** (e.g., increased battery capacity in updated EV models). To address this gap, I plan to explore high-dimensional network representations based on multilayer network theory [5]. In this approach, each layer will represent a time snapshot of the system's structure, and links between layers can track temporal changes in individual attributes. Additionally, I aim to develop a new network motif mining algorithm to identify statistically significant evolutionary patterns within this dynamic multilayer network. These patterns will then be inputs of a newly developed dynamic GNN to predict the future evolution of STSs. The **broader impact** of this project includes: 1) Beyond STSs, this new temporal modeling approach will also benefit other complex systems with temporal dynamics, such as social networks and protein-protein interactions (PPIs); 2) The development of the new network motif mining algorithm and dynamic GNN will advance the field of network science, particularly in the dimension of time.

2. AI-Integrated STS Engineering and Design

As AI increasingly permeates our daily lives, its interactions with both humans and machines are reshaping the socio-technical landscape. In smart manufacturing, for example, smart sensors monitor human health, while robots collaborate with humans on tasks like precision welding**. Although existing studies focus on individual-level human-AI interactions, a research gap remains in understanding and designing AI at higher levels—such as the subsystem (e.g., shop floor) and system-wide (e.g., entire factory) levels.** Addressing this gap requires tackling two key challenges: 1) developing methods to model complex physical systems that encompass diverse human, AI, and machine roles and their interactions, and 2) creating approaches to incorporate system-level interaction data into AI design, making AIs more adaptable to intricate environments. To meet the first challenge, I will model AI-integrated socio-technical systems as multi-agent systems (MAS) [6] and use complex networks to capture interactions among human, AI, and machine agents. This approach will include methods to quantitatively assess AI's impact on overall system performance. For the second challenge, I will apply GNNs to aggregate system-level interaction data and develop a general alignment projector that connects GNN outputs with the algorithms behind different AIs. Then, an optimization problem based on the combined model will be formulated and solved to improve AI adaptability and effectiveness. The **broader impact** of this project includes benefiting any AI-driven system (e.g., AI in Healthcare systems, etc.) by offering direct optimization strategies, while also providing AI developers with a system-level understanding of AI functionalities within complex socio-technical environments.

3. Exploring the Synergies between EV Market, Charging Infrastructure, and Power Grid

The **EV market**, **EV charging infrastructure**, and the **power grid** are three key components of the broader EV ecosystem. Significant efforts have been made separately to improve each of these systems. However, they are closely interconnected and impact one another. For example, EV sales directly influence the demand for charging stations, and the installation of charging stations affects the load on the power grid. Conversely, the capacity of the power grid can be a major constraint when deciding where to install new charging stations. In fact, according to our recent survey of EV customers, the availability of charging infrastructure is a major concern when deciding whether to purchase an EV. Given the socio-technical nature of these interconnected systems, along with my strong collaboration with industry (including research projects with Ford and an internship at Ford) and access to first-hand EV customer survey data, I am uniquely positioned to explore these systems in a more comprehensive and realistic way. Therefore, I propose a research project that not only analyzes the individual characteristics of these three systems using methods from my current work but also focuses on the interactions between them. For example, I plan to use machine learning techniques to develop regional customer profiles based on survey data, allowing for more accurate forecasting of charging demand. Additionally, I aim to develop a multilayer network model to analyze the interplay between EV charging infrastructure and the power grid. The **broader impact** of this project is twofold: 1) The research outcomes will directly contribute to the **Federal Government's 2030 goal of EV and EV charging**[2]. 2) Furthermore, I plan to integrate the analysis of these three systems into my teaching curriculum on network-based complex system engineering and design, helping to train the next generation of engineers specializing in EV ecosystems to support the national electrification plan.

Potential Collaborations and Funding Sources: One promising avenue for funding is the National Science Foundation (NSF), particularly its Operations and Design (OD) cluster programs within the Division of Civil, Mechanical, and Manufacturing Innovation (CMMI). These programs provide critical support for fundamental research in the design, operation, optimization, and control of complex engineered and socio-technical systems, which are integral to society's advancement. Additionally, relevant programs such as the Cyber-Physical System (CPS) in the CMMI division may offer additional opportunities. Moreover, given that socio-technical systems have broad applications across domains including, but not limited to, urbanization, transportation, energy, economics, and manufacturing, I also plan to seek research funding support from agencies such as the Defense Advanced Research Projects Agency (DARPA) (e.g., Complex Adaptive System Composition And Design Environment (CASCADE) program), Department of Transportation (DOT) (e.g., Strengthening Mobility and Revolutionizing Transportation (SMART) Grants Program), and the Department of Energy (DOE) (e.g., Smart Grid Investment Matching Grant Program). I am confident that the alignment of my research program with the interests of these government agencies in socio-technical systems and related areas will ensure its vitality and long-term sustainability. In addition, I will actively pursue collaborative opportunities with industry partners relevant to the above applications, including continuing my established collaboration with Ford on the evolutionary analysis of the U.S. EV market system. I also look forward to collaborating with department colleagues on research topics related to these fields.

References

[1] Y. Xiao, F. Ahmed, and Z. Sha, "Graph Neural Network-Based Design Decision Support for Shared Mobility Systems," *J. Mech. Des.*, vol. 145, no. 091703, Jun. 2023, doi: 10.1115/1.4062666.

[2] Y. Xiao, Y. Cui, M. T. Cardone, W. Chen, and Z. Sha, "Product Competition Analysis for Engineering Design: A Network Mining Approach," in *The Proceedings of the 2023 Conference on Systems Engineering Research*, D. Verma, A. M. Madni, S. Hoffenson, and L. Xiao, Eds., Cham: Springer Nature Switzerland, 2024, pp. 327–340. doi: 10.1007/978-3-031-49179-5_22.

[3] Y. Xiao and Z. Sha, "Robust design of complex socio-technical systems against seasonal effects: a network motif-based approach," *Des. Sci.*, vol. 8, p. e2, Jan. 2022, doi: 10.1017/dsj.2021.27.

[4] Y. Xiao, Y. Cui, J. Koskinen, N. Contractor, W. Chen, and Z. Sha, "Product Design Incorporating Competition Relations: A Network-Based Design Framework Considering Local Dependencies," *J. Mech. Des.*, vol. 147, no. 021702, Sep. 2024, doi: 10.1115/1.4066426.

[5] G. Bianconi, Ed., "Copyright Page," in *Multilayer Networks: Structure and Function*, Oxford University Press, 2018, p. 0. doi: 10.1093/oso/9780198753919.002.0003.

[6] P. Leitão, "Agent-based distributed manufacturing control: A state-of-the-art survey," *Eng. Appl. Artif. Intell.*, vol. 22, no. 7, pp. 979–991, Oct. 2009, doi: 10.1016/j.engappai.2008.09.005.

[2] https://www.whitehouse.gov/briefing-room/statements-releases/2023/04/17/fact-sheet-biden-harris-administration-announces-new-private-and-public-sector-investments-for-affordable-electric-vehicles/