Modeling Smart Cities with Hetero-functional Graph Theory

Wester C.H. Schoonenberg
Thayer School of Engineering
at Dartmouth College
Hanover, New Hampshire
Wester.C.H.Schoonenberg.TH@Dartmouth.edu

Amro M. Farid
Thayer School of Engineering
at Dartmouth College
Hanover, New Hampshire
Email: amfarid@dartmouth.edu

Abstract—In the 21st century, urbanization as a mega-trend will create many megacities. These highly dense, large population centers will have to efficiently deliver essential services including, energy, water, mobility, manufactured goods, and healthcare. While these services may be treated independently, they are in reality interdependent, especially as the need for efficient resource utilization, and consequently integration. This presents a formidable engineering challenge as the modeling foundations for these services have traditionally been discipline specific. Furthermore, efforts to integrate these modeling foundations have often adopted simplifying constraints which have limited applicability to the emerging challenge of smart cities. This paper collates an emerging “hetero-functional graph theory” for potential application to integrated smart city infrastructure models. It has been recently demonstrated in several application domains. The paper concludes with the construction of a hetero-functional graph for a smart city model consisting of an integrated electricity, water, and transportation system. Such a graph has the potential for dynamic modeling, resilience analysis, and integrated decision-making.

Keywords – Hetero-functional Graph Theory, Infrastructure, Smart City, Axiomatic Design, Engineering Systems

I. INTRODUCTION

In the past 50 years, the percentage of global population living in cities has increased from 34% to 52% [1]. This massive migration has resulted in high density urban areas and increases the stress on the existing infrastructure [2]. The existing infrastructure delivers essential services including energy, water, mobility, manufactured goods, and healthcare. These infrastructures are not only complex and of a large size, but also interdependent, especially as there is need for efficient resource utilization, and thus integration. The interactions in this “system of systems” create unpredictable and poorly understood behavior.

Other trends in urban environments, are to include technology and improve efficiency based on the data generated by the technology. Examples include smart meters, smart thermostats, smart lighting [3], but also smart traffic control [4]. In order to understand the measured behavior of the complex systems, a model needs to be developed that includes the interdependencies amongst the networks. For network based systems, such as transportation systems, power grids, water networks, supply chains, and healthcare systems, a graph theoretic approach has formed the basis of much research [5]. Examples include the maximization of power system resiliency [6]–[8], shortest path calculations [9], hospital operations [10], supply chain management [11], and even brain function [12]. Graph theory offers many advantages, such as proven optimal solutions for certain types of problems [13]. Disadvantages of graph theory include the limited capability to capture more than one discipline with the existing mathematics.

The network sciences community has made effort to introduce the multilayer features to improve the understanding of complex systems [14]. This approach has the potential to facilitate modeling of “systems of systems” that cross disciplinary boundaries. However, the paper written by Kivela also discusses the constraints that are imposed by existing literature [14]. These constraints limit the potential application domains of multilayer network theory. For a smart city model as described in this paper, a different method needs to be introduced.

Recently, a graph theoretic modeling approach called “Hetero-functional Graph Theory” has developed from roots in the Axiomatic Design for Large Flexible Engineering Systems. It provides a rigorous platform for modeling systems of systems [15]. An essential aspect of this platform is that it defines mutually exclusive and collectively exhaustive sets of system processes and resources. In so doing, it facilitates the translation of systems engineering models (e.g. SysML) to a mathematical engineering systems description. Hetero-functional Graph Theory has been applied in transportation [16], healthcare [17], production systems [18], power grid [19], and resilience studies [15]. Additionally, hetero-functional graph theory has proven to be able to support modeling of two unlike systems. This paper applies hetero-functional graph theory to model a smart city test case in a single mathematical model.

The basis for Hetero-functional Graph Theory is introduced in Section II. Section III introduces Hetero-functional Graph Theory for individual infrastructure systems. The combination of these systems form a Smart City in Section IV. Section V concludes the paper and provides an outlook to future work.

II. HETERO-FUNCTIONAL GRAPH THEORY

This section succinctly covers the need for Hetero-functional Graph Theory. It then describes its foundations as
a modeling approach for multi-layer engineering systems.

A. Need for Hetero-functional Graph Theory

The emerging challenge of smart cities calls for modeling frameworks that cross disciplinary boundaries. In recent years, the network sciences community has developed significant work in “multi-layer networks” [14] which has the potential to cross disciplinary boundaries. Nevertheless, a recent review on multi-layer networks has identified that many of these models adopt simplifying constraints. The discussed methods impose one or more of the following constraints [14]:

1) Alignment of nodes amongst the layers.
2) Requirement that the layers need to be disjoint.
3) All layers have the same number of nodes.
4) The couplings are diagonal.
5) The inter-layer couplings consist of layer couplings.
6) The inter-layer couplings are categorical.

Such constraints likely limit the number of practical applications, such as modeling smart cities. These integrative models should have the ability to incorporate arbitrary couplings amongst layers. In that regard, this work collates an emerging “hetero-functional graph theory” for potential application to smart cities.

B. Foundations of Hetero-functional Graph Theory

This subsection introduces the very elemental basics of hetero-functional graph theory, a more in depth discussion can be found in the provided references. The foundations for hetero-functional graph theory are ontological. Ontological science encourages the development of models that have the properties of soundness, completeness, lucidity, and laconicity, which are conceptually understood from Figure 1 [20]. Ultimately, these four properties taken together require a 1-to-1 mapping of conceptualized abstractions and formal mathematical models [21]. In this work, hetero-functional graph theory is based upon three abstract concepts which may be found in the systems engineering and engineering design literature: (1) Allocated Architecture, (2) Hetero-functional Adjacency Matrix, and (3) Controller Agency Matrix. Each of these is now addressed in turn.

1) Allocated Architecture: The allocated architecture is a structural description of a system’s capabilities. Mathematically, the allocation of system processes \( P \) to resources \( R \) is described by a “design equation” [15], [21]–[23]:

\[
P = J_S \odot R
\]

where \( J_S \) is a binary matrix called a “system knowledge base”, and \( \odot \) is “matrix boolean multiplication” [15], [21]–[23]. The design equation originates from Axiomatic Design for Large Flexible Engineering Systems and requires that these system processes and resources be mutually exclusive and collectively exhaustive [15], [22], [23]. The system resources are a combination of transformation, buffering, and transporting resources, and the system processes are a combination of transformation, holding, and transporting processes. The knowledge base \( J_S \) can thus be constructed from the mappings of these subcategories, where \( J_M \) is the transformation knowledge base, and \( J_H \) the refined transportation knowledge base. Their combinations are provided in Equation 2, and the constraints matrix is constructed in a similar way (Equation 3) [15], [22], [23].

\[
J_S = \begin{bmatrix} J_M & 0 \\ J_H & 0 \end{bmatrix}
\]

\[
K_S = \begin{bmatrix} K_M & 0 \\ K_H & 0 \end{bmatrix}
\]

Definition 1. System Knowledge Base [15], [22], [23]: A binary matrix \( J_S \) of size \( \sigma(P) \times \sigma(R) \) whose element \( J_S(w,v) \in \{0,1\} \) is equal to one when action \( e_{w,v} \in \mathcal{E} \) (in the SysML sense) exists as a system process \( p_w \in P \) being executed by a resource \( r_v \in R \). The \( \sigma() \) notation gives the size of a set.

The system knowledge base itself forms a bipartite graph and has been defined for several different disciplines, such as transportation [16], power [19], mass-customized production [24], water distribution, and health-care delivery systems [17].

The system’s capabilities are quantified as structural degrees of freedom [22], [23]. The existence of these capabilities are represented by filled elements in the system knowledge base. Meanwhile, the constraint matrix describes the availability of these capabilities in recognition that many systems change their capabilities over time.

Definition 2. System Events Constraints Matrix [15], [22], [23]: A binary matrix \( K_S \) of size \( \sigma(P) \times \sigma(R) \) whose element \( K_S(w,v) \in \{0,1\} \) is equal to one when a constraint eliminates event \( e_{w,v} \) from the event set.

Definition 3. Structural Degrees of Freedom [15], [22], [23]: The set of independent actions \( \psi_i \in \mathcal{E}_S \) that completely
defines the available processes in a large flexible engineering system. Their number is given by:

\[
DOP_S = \sigma(\mathcal{E}_S) = \sum_w \sum_v (J_S \odot K_S)(w,v) \tag{4}
\]

where the potential for interconnected capabilities is described by the system sequence knowledge base \(J_p\) and the availability of these connections is described by the system sequence constraints matrix \(K_p\).

**Definition 4. Sequence-Dependent Degrees of Freedom**

[15], [25], [26]: The set of independent pairs of actions \(z_{\psi_1\psi_2} = e_{w_1v_1}e_{w_2v_2} \in Z\) of length 2 that completely describe the system language. The number is given by:

\[
DOP_p = \sigma(Z) = \sum_{\psi_1} \sum_{\psi_2} (J_p \odot K_p)(\psi_1, \psi_2) \tag{7}
\]

where \(J_p\) and \(K_p\) as defined in Definition 5 and 6.

**Definition 5. System Sequence Knowledge Base**

[15], [25], [26]: A square binary matrix \(J_p\) of size \(\sigma(P)\sigma(R)\times\sigma(P)\sigma(R)\) whose element \(J_p(\psi_1, \psi_2)\) is equal to one when string \(z_{\psi_1\psi_2}\) exists. It may be calculated directly as

\[
J_p = [J_S \cdot K_S]^V [J_S \cdot K_S]^{VT} \tag{9}
\]

where \((\cdot)^V\) is shorthand for vectorization (i.e. vec()).

**Definition 6. System Sequence Constraints Matrix**

[15], [25], [26]: a square binary constraints matrix \(K_p\) of size \(\sigma(P)\sigma(R)\times\sigma(P)\sigma(R)\) whose elements \(K_p(\psi_1, \psi_2)\) are equal to one when string \(z_{\psi_1\psi_2} = e_{w_1v_1}e_{w_2v_2} \in Z\) is eliminated.
and the holding processes $P_{IE}$ are “carry power at 132 kV”. The mapping of processes and resources is achieved with the knowledge base $J_{SE}$, following Definition 1. This knowledge base is the result of a combination of the transformation knowledge base $J_{ME}$ and the refined transportation knowledge base $J_{HE}$, as demonstrated in general in Equation 2. The corresponding constraints matrix $K_{SE}$ is similarly constructed in general in Equation 3.

B. Transportation Systems

Previous work has modeled transportation systems with the use of Hetero-functional Graph Theory [26]. The transportation system resources are defined as: $R_T = M_T \cup B_T \cup H_T$, where $M_T$ are the transformation resources (e.g. stations that enter people in and out of the system), $B_T$ are the independent buffers (e.g. parking lots), and $H_T$ the transporters (e.g. roads). The transportation system processes are defined as: $P_T = P_{\mu T} \cup P_{\gamma T} \cup P_{\eta T}$, where $P_{\mu T}$ are the transformation processes (e.g. enter the transportation system), $P_{\gamma T}$ are the transportation processes (e.g. transport a person form node 1 to node 3), and $P_{\eta T}$ are the holding/charging processes (e.g. charge the vehicle). The transportation system knowledge base and the constraints matrix can be constructed using Equations 2 and 3.

IV. HFGT Infrastructure Model for Smart City Infrastructures

A. Case study: power, transportation, and water system

The smart city case study is based on the Symmetrica case study, which has been used in previous work in transportation-electrification [16]. The Symmetrica study combines a 201-bus power system (Figure 2), a symmetric 13×13 node transportation system (Figure 3), and a 125 node potable water system (Figure 4). The data for the test case is openly available [28].

1) Power System: The Symmetrica power grid lay-out (Figure 2) is derived from the 201-bus distribution system test case [29], [30]. The transformation resources $M_E$ are a generator at Node 201, and consuming resources on all other nodes. Some of these consumption nodes are coupled to the transportation and water distribution systems. The transportation resources are the power lines as indicated in the figure.

2) Transportation System: The Symmetrica transportation system lay-out (Figure 3) was defined in previous work as a test case for transportation studies, comparable to the test cases used in power systems [27]. The transport network consists

C. Water Distribution Systems

Previous work has modeled water distribution systems with the use of Hetero-functional Graph Theory [15]. The water system resources are defined as $R_W = M_W \cup B_W \cup H_W$, where $M_W$ are the transformation resources (e.g. a water treatment facility), $B_W$ are the independent buffers (e.g. pipe junctions), and $H_W$ are the transporters (e.g. water pipes). The water system processes are defined as $P_W = P_{\mu W} \cup P_{\eta W} \cup P_{\gamma W}$, where $P_{\mu W}$ are the transformation processes (e.g. generate water), $P_{\eta W}$ are the transportation processes (e.g. transport water from junction 2 to junction 13), and $P_{\gamma W}$ are the holding processes (e.g. pressurize water). Similar to the previous systems, the water distribution system knowledge base and constraints matrix are constructed in Equations 2 and 3.
of 169 nodes; all of which are independent buffers \((B_T)\). The holding processes differentiate the buffers to allow for charging at 53 charging nodes. The transportation resources are the roads between the nodes, but notice that each road in the figure represents two transportation resources, as the roads are bidirectional and allow a two way flow of traffic. Some roads are enabled to wirelessly charge the driving vehicles, this is again achieved by including a holding process at the 104 electrified roads.

3) Water Network: The water network is based on the Anytown Water Distribution Network lay-out [31]. The network is adjusted and now consists of 5 water treatment facilities, 10 water storage tanks, and 110 water consumption nodes (Figure 4). The treatment facilities and the consumption nodes are the transformation resources \(M_W\), and the storage tanks are independent buffers \(B_W\). The pipes are the transporting resources \(H_W\). Note that each of the nodes in the system is connected to the electric grid, as they all include electric water pumps.

Collectively, the nodes and edges in each infrastructure constitute the system resources. The legend in Figure 2 contains the couplings between the power grid and the other two networks. Care must be taken to avoid the double counting of resources which exist in multiple networks. For example, node 71 in the power grid, node 1 in the transportation network and node 20 in the water network effectively represent one city resource (e.g. building). Consequently, in the knowledge base for the entire city, resource 71 contains capabilities for power consumption, vehicle charging, and water consumption. Also note that all nodes in the water network consume power, in this situation there is a one-to-one mapping from water nodes onto power nodes. The transportation knowledge bases are refined individually and aggregated afterwards to maintain the definition of holding processes. The number of capabilities in the system as defined in Equation 10 is 2,248. The capabilities are projected as nodes in a three dimensional space in Figure 5. The top layer is the water system, the middle layer is the power grid, and the bottom layer is the transportation system.

4) Integration of Multiple Infrastructure Systems: Based upon the figures of the individual infrastructures, and their associated data [28], the knowledge base for the multiple infrastructure systems can be constructed.

Collectively, the nodes and edges in each infrastructure constitute the system resources. The legend in Figure 2 contains the couplings between the power grid and the other two networks. Care must be taken to avoid the double counting of resources which exist in multiple networks. For example, node 71 in the power grid, node 1 in the transportation network and node 20 in the water network effectively represent one city resource (e.g. building). Consequently, in the knowledge base for the entire city, resource 71 contains capabilities for power consumption, vehicle charging, and water consumption. Also note that all nodes in the water network consume power, in this situation there is a one-to-one mapping from water nodes onto power nodes. The transportation knowledge bases are refined individually and aggregated afterwards to maintain the definition of holding processes. The number of capabilities in the system as defined in Equation 10 is 2,248. The capabilities are projected as nodes in a three dimensional space in Figure 5. The top layer is the water system, the middle layer is the power grid, and the bottom layer is the transportation system.

![Figure 4: Water Network for Symmetrica [28]](image)

4) Integration of Multiple Infrastructure Systems: Based upon the figures of the individual infrastructures, and their associated data [28], the knowledge base for the multiple infrastructure systems can be constructed.

Collectively, the nodes and edges in each infrastructure constitute the system resources. The legend in Figure 2 contains the couplings between the power grid and the other two networks. Care must be taken to avoid the double counting of resources which exist in multiple networks. For example, node 71 in the power grid, node 1 in the transportation network and node 20 in the water network effectively represent one city resource (e.g. building). Consequently, in the knowledge base for the entire city, resource 71 contains capabilities for power consumption, vehicle charging, and water consumption. Also note that all nodes in the water network consume power, in this situation there is a one-to-one mapping from water nodes onto power nodes. The transportation knowledge bases are refined individually and aggregated afterwards to maintain the definition of holding processes. The number of capabilities in the system as defined in Equation 10 is 2,248. The capabilities are projected as nodes in a three dimensional space in Figure 5. The top layer is the water system, the middle layer is the power grid, and the bottom layer is the transportation system.

![Figure 5: Representation of the Smart City Degrees of Freedom, with layer specific sequence dependent degrees of freedom](image)

B. Network Visualization

After constructing the knowledge base for the full city, the allocated architecture is now used to form the basis for the calculation of the hetero-functional adjacency matrix and the corresponding sequence dependent degrees of freedom. For clarity, they are presented in Figure 5 as the connections between the nodes for each layer. Figure 6 also contains the sequence dependent degrees of freedom in between the layers. Their total number is 21,145.

V. CONCLUSION AND FUTURE WORK

This paper introduces hetero-functional graph theory as a method to model multi-disciplinary engineering systems. Hetero-functional graph theory allows for the integration of layers of networks without the constraints imposed by previous methods. The paper applies hetero-functional graph theory
to create a three layer smart city infrastructure model. In the model, function and form are represented unambiguously. Finally, the paper shows a visual representation of all system capabilities, and includes the allowed sequences of these capabilities.

Future work includes the expansion of the model to include more infrastructures. Additionally, these hetero-functional graphs may form the basis for measuring multi-layer infrastructure resilience.

REFERENCES


