A GENERIC FLASHOVER PREDICTION MODEL FOR RESIDENTIAL BUILDINGS USING GRAPH NEURAL NETWORK

Wai Cheong Tam¹,4, Eugene Yujun Fu²,4, Paul Reneke³, Richard Peacock⁴, Thomas Cleary⁵

ABSTRACT

A generic graph neural network-based model is developed to predict the potential occurrence of flashover for different building structures. The proposed model transforms multivariate temperature data into graph-structure data. Utilizing graph convolution operations, the temporal dependencies and spatial correlations of the temperature data are captured. Model assessment show that the generic flashover prediction model can distinguish different building structures and provide forecasts in advance to classify the potential occurrence of flashover with an overall accuracy of ~ 93 %. This work constitutes a machine learning-based forecasting model framework accounting for a wide range of building structures. The research outcomes from this study are expected to facilitate data-driven fire fighting, leading to enhanced situational awareness and improved fire fighting safety to help reduce U.S. fire fighter deaths and injuries.

Keywords: Graph neural network; flashover prediction; compartment fires; synthetic fire data

1 INTRODUCTION

In the United States, fire fighter deaths and injuries are a concern. In the years from 2008 to 2018, statistics show that approximately 800 fire fighters have been killed and more than 320,000 are injured [1,2]. Based on a recent study conducted by the National Fire Protection Association (NFPA) in 2020 [1], approximately 58 % of fire ground deaths occur in residential fires where about 46 % of deaths are related to fires in one- or two-family homes (~ 23 %) and apartments (~ 23 %). The remaining 8 % of deaths are in unspecified homes. It is worth noting that many injuries are career ending type of injuries. In the same studies [1,2], it is revealed that rapid fire progression, such as flashover, has been one of the major causes of fire fighter deaths (~ 13%) and injuries. However, fire fighters have not had any tools to predict flashovers and rely solely on their experience to recognize flashover indicators [3], such as rollover near ceiling and/or high heat, to avoid this life-threatening event. Yet, since the fire conditions change rapidly and the transition of flashover usually happens within seconds [3], it is rather difficult to identify the imminent occurrence of flashover while inside a building fire and hardly possible on the outside. With that, if the fire fighters cannot recognize impending flashover their lives are at tremendous risk. It is believed that with the advancement of smart interconnected fire protection devices and systems [4], a data-driven model can be developed to

---

¹ Mechanical Engineer, Fire Research Division, National Institute of Standards and Technical, Gaithersburg, Maryland, USA e-mail: waicheong.tam@nist.gov, ORCID: https://orcid.org/0000-0003-3913-0243
² Research Assistant Professor, Department of Rehabilitation Sciences, The Hong Kong Polytechnic University, Hung Hom, Hong Kong SAR e-mail: eugene.fu@polyu.edu.hk, ORCID: https://orcid.org/0000-0003-1048-1904
³ Computer Scientist, Fire Research Division, National Institute of Standards and Technical, Gaithersburg, Maryland, USA e-mail: paul.reneke@nist.gov
⁴ Chemical Engineer, Fire Research Division, National Institute of Standards and Technical, Gaithersburg, Maryland, USA e-mail: richard.peacock@nist.gov
⁵ Chemical Engineer, Fire Research Division, National Institute of Standards and Technical, Gaithersburg, Maryland, USA e-mail: thomas.cleary@nist.gov, ORCID: https://orcid.org/0000-0001-8785-2035

Joint first authors
predict the potential occurrence of flashover based on typical thermal sensor signals, such as from heat detectors.

Over the past twenty years, many attempts have been made to develop prediction models that can be used in building fires. Existing prediction models are based on empirical correlations [5-7], inverse modelling techniques [8,9], and sensor-steered CFD approaches [10,11]. Yet, the models [5,8] are generally limited to single-compartment applications and are not suitable for hazard predictions in multi-compartment structures. Although the sensor-steered CFD approaches [10,11] provide much better prediction capabilities and higher accuracy, the drawbacks are that they rely on the use of high-performance computing machines and require relatively long computational time. In [10], it is stated that it will take approximately five minutes to provide a reliable prediction for one-time step. Therefore, models from [10,11] are impractical to implement in real-time fire fighting applications.

Another limitation associated with existing models [5-11] is that they rely on both the continuous temperature signals from thermocouples in all compartments and the prior knowledge about the fire locations and vent opening conditions. Yet, fire protection devices, such as heat detectors, will stop functioning at elevated temperature [12] and the information about the fire locations and the opening conditions of windows and doors are generally unknown. Since these realistic conditions (i.e., sensor temperature limit and the effect of arbitrarily fire location and vent opening conditions) have not been considered in the development process, the model performance from [5-11] are likely to be diminished. A more robust approach is needed to overcome both the numerical challenge and the complexity of realistic conditions.

Machine learning (ML) paradigms can provide a breakthrough to problems in various scientific and engineering communities. Examples can be found such as fault detection [12], cyber-attack prevention [13], object detection for self-driving vehicles [14], etc., in which real time predictions with reliable accuracy are needed. In the fire research community, efforts have also been made to develop ML-based models to tackle complex problems [15-17]. For fire fighting in residential buildings, recent works [18,19] provide practical solutions utilizing ML paradigms to predict the potential occurrence of flashover with lead time based on limited sensor temperatures up to 150 °C or 250 °C. In [19], the model performance is also verified with real fire experiments and it is demonstrated that the model can predict the potential occurrence of flashover 30 seconds in advance with an overall accuracy of ~ 82 %. It is also highlighted that the model takes less than a second to provide a prediction. In comparison to [10,11], the numerical improvement is substantial. However, the prediction models from [18,19] are only developed for a specified building structure and will not work in other building structures. Also, the model structure will not be able to handle different numbers of inputs (i.e., four channels of temperature data in one building structure vs six channels of temperature data in another building structure). In principle, one can argue that additional models can be developed for other building structures, but this brute force approach is not efficient for a community-based system. Therefore, a generic flashover prediction developed based on a graph neural network is proposed to encompass a range of building configurations. The model will provide prediction of potential flashover occurrence in the future using only the limited available temperature without any prior knowledge about the layout of the building structure and information about the fire (i.e., location, size, growth) and opening conditions of doors and windows. In the following sections, the data collection process is described and the model formulation is presented. Results and discussion are given to demonstrate the effectiveness of the model.

2 COLLECTION OF DATA

The development of a robust ML-based flashover prediction model requires a dataset that adequately covers the intended application domain. In principle, since flashover is an extreme fire event, the acquisition of data is a challenging task. In a recent study [16], approximately 1000 full-scale experiments were needed to develop an accurate prediction model for an idealized three-compartment building structure with simplified fire and vent opening conditions. For this study in which a significant number of building structures will have to be accounted for, the required number of experiments increases proportionally to the
number of building structures, the total number of compartments within a building structure, fire locations, fire growth rates, and the vent opening conditions. It can easily be understood that physically conducting such a huge set of full-scale experiments is not feasible. With that, a learning-by-synthesis approach is used to obtain the fire data needed to facilitate the model development.

2.1 CData (CFAST Fire Data Generator)
The learning-by-synthesis (LBS) approach [20] refers to the use of computer simulation programs to carry out the desired full-scale experiments numerically. Specifically, a zone model, namely CFAST [21], is utilized here. CFAST is a two-zone fire model that predicts the thermal environment caused by a fire within a compartmented structure. The model is mathematically verified and validated against more than 15 sets of full-scale experiments with a wide range of different parameters [22]. In terms of accuracy, CFAST predictions of upper layer gas temperature average within 6 % of experimental measurements.

Although CFAST can be used to carry out fire experiments numerically, manually creating thousands of input files is a time-consuming task. In order to carry out the data collection process systematically in this study, CData [23] is utilized to automatically generate the desired number of CFAST input files. As mentioned in [23], CData is a CFAST program that can operate on a PC computer desktop or a Linux cluster. Given the user-specified simulation parameters, such as building layouts, surface materials, fire conditions, ventilation configurations, location of detector(s), and output intervals, the program provides corresponding CFAST input files based on seven different probability density functions. User-defined functions can also be used to formulate inputs for desired fire environments. In this study, approximately 230 nodes were used to generate the fire data and the data generation process (excluding the time required to setup the input files and data preprocessing) was about two weeks. A summary spreadsheet containing detailed information about all input parameters for each of the cases is also generated. This spreadsheet can be used to carry out data inspection to eliminate any duplicate cases and to examine the data behavior of each of the simulation parameters. This information is crucial to ensure the quality and the covering range of the fire data.

2.2 Home selection
The focus of this study was on residential structures, and seventeen homes are selected for the development of a generic flashover prediction model. These homes are selected from [24] in which 209 dwellings are defined to provide a representation of approximately 80% of the U.S. housing stock. In this study, the seventeen homes being considered are all single-floor building structures and they are categorized into three types: 1) detached homes, 2) attached homes, and 3) apartment homes. The overall floor area is from 65 m² to 275 m² with the total number of compartments ranging from three to fourteen. Three examples are provided in Figs. 1 with Fig. 1a, 1b, and 1c being an apartment home (#2), attached home (#9), and a detached home (#17), respectively. The complete list of the seventeen homes is provided in Appendix A for ease of reference. As seen in Fig. 1a, the apartment home consists of five different compartments including a living room, kitchen, bedroom, bathroom, and a hallway. In Fig. 1c, the hallway is divided into three sections and this process is done to facilitate the use of CFAST for data generation. Table 1 provides the overall floor area and interior details of the seventeen homes. It should be noted that the current set of seventeen homes were carefully selected to include two modelling challenges: 1) buildings that have identical number of compartments with a small compartment arrangement in the building layout (see home #4 and home #5 in Appendix A) and 2) buildings that are completely different with both the total number of compartments and the layout itself. For that, the proposed model will need to correctly identify the building structure based on the learned compartment relationships from the time-series input data. Also, it is worth noting that one can create the numerical models for all 209 dwellings to develop a more complete generic model. This research effort in on-going and it will be reported in a future publication.

2.3 Numerical setup
The numerical setup of all models is relatively similar for the seventeen selected homes. The overall ceiling height is 2.4 m. For all the doors and doorways, the height and the width are 2.05 m and 0.9 m, respectively.
Figures 1. An example of the building layout for a) a five-compartment apartment home (#2), b) a six-compartment attached home (#9), and c) a fourteen-compartment detached home (#17).

Table 1. Housing details for the seventeen selected homes.

<table>
<thead>
<tr>
<th>Home #</th>
<th>Type*</th>
<th>Total Compartment</th>
<th>Floor Area (m²)</th>
<th>Living Room</th>
<th>Kitchen</th>
<th>Hall†</th>
<th>Bed-Room</th>
<th>Bath-Room</th>
<th>Dining Room</th>
<th>Family Room</th>
<th>Den</th>
<th>Fire Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>APT</td>
<td>3</td>
<td>65</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2000</td>
</tr>
<tr>
<td>2</td>
<td>APT</td>
<td>5</td>
<td>65</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3000</td>
</tr>
<tr>
<td>3</td>
<td>APT</td>
<td>6</td>
<td>65</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4000</td>
</tr>
<tr>
<td>4</td>
<td>APT</td>
<td>7</td>
<td>65</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4000</td>
</tr>
<tr>
<td>5</td>
<td>APT</td>
<td>6</td>
<td>65</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4000</td>
</tr>
<tr>
<td>6</td>
<td>APT</td>
<td>7</td>
<td>65</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4000</td>
</tr>
<tr>
<td>7</td>
<td>APT</td>
<td>8</td>
<td>65</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>5000</td>
</tr>
<tr>
<td>8</td>
<td>AH</td>
<td>5</td>
<td>95</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3000</td>
</tr>
<tr>
<td>9</td>
<td>AH</td>
<td>6</td>
<td>95</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4000</td>
</tr>
<tr>
<td>10</td>
<td>AH</td>
<td>7</td>
<td>95</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>5000</td>
</tr>
<tr>
<td>11</td>
<td>DH</td>
<td>7</td>
<td>107</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5000</td>
</tr>
<tr>
<td>12</td>
<td>DH</td>
<td>7</td>
<td>107</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>5000</td>
</tr>
<tr>
<td>13</td>
<td>DH</td>
<td>8</td>
<td>107</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>6000</td>
</tr>
<tr>
<td>14</td>
<td>APT</td>
<td>8</td>
<td>130</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>4000</td>
</tr>
<tr>
<td>15</td>
<td>APT</td>
<td>8</td>
<td>142</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>5000</td>
</tr>
<tr>
<td>16</td>
<td>DH</td>
<td>13</td>
<td>275</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>4</td>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>8000</td>
</tr>
<tr>
<td>17</td>
<td>DH</td>
<td>14</td>
<td>275</td>
<td>1</td>
<td>1</td>
<td>3</td>
<td>3</td>
<td>3</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>8000</td>
</tr>
</tbody>
</table>

* Note that apartment home, attached home, and detached home are denoted as APT, AH, and DH, respectively. † No fire in these compartments. In total, there are 78000 cases for the seventeen different structures.
The horizontal location of windows (shown as blue lines) is centred in a compartment wall as seen in Fig. 1. The width, height, and vertical location of the windows are 1.8 m x 1.4 m x 0.6 m for the living room, 1.4 m x 0.85 m x 1.2 m for the kitchen, 1.4 m x 1.4 m x 0.6 m for the dining room, 1.8 m x 1.4 m x 0.6 m for the family room, and 1.4 m x 1.4 m x 0.6 m for the bedrooms. The glazing material is consistent for all windows, 3 mm single-pane float glass [22]. The thermal properties are shown in Table 2. In addition, there is a heat detector in every building compartment. They are generally located near doors or doorways and are placed about 0.02 m below from the ceiling. The locations of the heat detectors are shown as red dot symbols in Fig. 1. The response time index for the heat detector is 35 (ms)^{0.5}. Temperature data obtained from the heat detectors are used for model development. In terms of interior finish, the walls and the ceiling of each compartment are constructed with gypsum wallboards and the floor is built with concrete. The thermal properties of the materials are provided in Table 2.

<table>
<thead>
<tr>
<th>Materials</th>
<th>Conductivity (kW/(m-K))</th>
<th>Specific Heat (kJ/(kg-K))</th>
<th>Density (kg/m^3)</th>
<th>Thickness (m)</th>
<th>Emissivity (-)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glazing</td>
<td>0.0008</td>
<td>0.8</td>
<td>2500</td>
<td>0.003</td>
<td>0.95</td>
</tr>
<tr>
<td>Gypsum</td>
<td>0.00016</td>
<td>1</td>
<td>480</td>
<td>0.025</td>
<td>0.9</td>
</tr>
<tr>
<td>Concrete</td>
<td>0.0016</td>
<td>0.75</td>
<td>2400</td>
<td>0.15</td>
<td>0.94</td>
</tr>
</tbody>
</table>

2.4 Fire conditions

One fire is initiated in one compartment for each numerical experiment. The fire location is at the center of any compartments, except the hallways and bathrooms. The fire growth behavior is divided into four different stages. As shown in Fig. 3, the stages include linear growth (smoldering fire), t-squared growth (flaming fire), peak, and decay. In this study, three furniture items are selected and they are chairs, polyurethane foam mattresses, and cotton-based spring mattresses. Table 3 specifies the fire growth parameters of the three items: the transition heat release rate (HRR) from smoldering fire to flaming fire (Q1), peak HRR (Q_{max}), time to transition (t1), time to peak HRR (t2), peak time (t3 – t4), and decay time (t4 – t2). The peak HRR and time to peak HRR are obtained from [23] and the fire growth rate is between 0.000329 kW/s^2 and 0.041387 kW/s^2, ranging from slow to fast fire growth rate. In general, the selection of these items follows that of the primary burning items in [24]. Transition HRR and time to transition HRR are selected based on examination of HRR data provided in [25]. Since the flashover usually occurs during the t-squared growth stage and peak stage, the exact value for peak time and decay time is not significant. One thousand different fire cases are assigned to each compartment and the total number of fire cases for each building structure is shown in Table 1.

![Figure 3. Overview of HRR in different fire growth stages for a typical t-squared fire.](image)

<table>
<thead>
<tr>
<th>Items</th>
<th>Q1 (kW)</th>
<th>Q_{max} (kW)</th>
<th>t1 (s)</th>
<th>t2 (s)</th>
<th>t3-t2 (s)</th>
<th>t4-t3 (s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chair</td>
<td>10 – 30</td>
<td>270 – 3500</td>
<td>150 – 1200</td>
<td>295 – 675</td>
<td>1000</td>
<td>100</td>
</tr>
<tr>
<td>Mattress (foam)</td>
<td>10 – 30</td>
<td>2275 – 4620</td>
<td>150 – 1200</td>
<td>305 – 435</td>
<td>1000</td>
<td>100</td>
</tr>
<tr>
<td>Mattress (cotton)</td>
<td>10 – 30</td>
<td>130 – 1670</td>
<td>150 – 1200</td>
<td>360 – 1240</td>
<td>1000</td>
<td>100</td>
</tr>
</tbody>
</table>
In CData, probability density functions (PDF) are used to facilitate case samplings and the use of a specific PDF is based on numerical experiments such that a wide range of fire growth behavior is included and sufficient data for flashover cases are obtained. Given the fact that a ML based model is intended to be developed for potential flashover prediction and since there will be a significant data imbalance associated with non-flashover related and flashover related data, a uniform PDF is used to sample peak HRR for the three different items and a skew normal PDF is used to sample time to peak HRR. The motivation is that more flashover cases (i.e., more rapidly growing fires) can be included in the data set so that the ML based model will determine the important relationship about potential flashover conditions. For transition HRR, time to transition, peak time, and decay time, a uniform PDF is utilized.

2.5 Vent opening conditions

There are three types of vents in this study: interior doors such as bedroom and bathroom doors, exterior doors such as front doors, and windows. For interior doors, the current setting assumes all interior doors to be opened from the beginning for all tests. This simplification is made to reduce the complexity of the data behavior, but this constraint will be removed in future studies. For front doors, there is a 50% chance that the door is opened at the beginning of a test and this will provide fresh air to sustain the fire growth. In terms of CData specifications, a uniform PDF is used to sample the opening door event. For windows, a temperature-trigger setting is used to allow the windows to be arbitrarily opened when a temperature set-point is reached. Based on [24], breakage of a single-pane float glass is experimentally observed at temperature in between 100 °C and 200 °C. An average value of 150 °C is used to be the temperature set-point. In order to adopt the window breakage phenomenon, a target is placed at the top of a window. The direction of the target is normal to the window surface and the target thermal properties are taken to be that of a 3 mm single-pane float glass as shown in Table 1. It is worth noting that although the current fire and vent opening conditions are relatively simple, the generated data do cover a wide range of realistic fire scenarios. Also, given the fact that data will have to be generated for seventeen different buildings, the data size is substantial. Provided the numerical setting mentioned above, there are a total of 78000 cases.

2.6 Data profiles and the effect of maximum operational temperature limit for detectors

Figure 4a shows six heat detector local temperature readings for a medium growth fire (~ 0.014 kW/s²) with high peak HRR (~ 3060 kW) occurring in the living room from an attached home as shown in Fig. 1b. The total simulation time for each numerical experiment is 3600 s and the temperature output interval is 5 s. In this case, all interior and front doors are always open. Given sufficient oxygen, the fire continues to grow and the detector temperature increases accordingly until the fire goes out. The temperature oscillation around 400 s is due to broken living room window. It should be noted that the label (red dash line) is the upper layer gas temperature for the room of fire origin and it is used to determine when the temperature threshold for flashover condition is met.

![Figures 4. Temperature profiles of heat detectors at different compartments a) without sensor limits and b) with a sensor limit at 150 °C for case #9 as shown in Fig.1b.](image-url)
A proper temperature criterion is carefully selected for flashover conditions. In [26], the onset of flashover temperature ranges from 450 °C to 771 °C. This wide range of values for both measurable criterions is due to the huge change of temperature during the transition to flashover. Yet, it is observed from [4] that most of the values are in the 550 °C to 650 °C range. To be conservative and following recommendations from [22], the upper gas layer temperature at 600 °C is used as the threshold for the potential occurrence of flashover in this study.

Loss of heat detector temperature signal is a realistic fire damage condition that adds another layer of complexity to the current problem. In actual fire scenarios, heat detectors are very unlikely to survive at elevated temperature [27] approaching flashover conditions and will fail at temperatures well below flashover occurrence. According to NFPA 72 [28], heat sensing fire detectors are categorized into seven different classes with temperature ranging from low to ultra-high and the maximum operational temperature ranging from 29 °C to 302 °C. In this study, a temperature cut-off of 150 °C is adapted. As an illustration, Fig. 4b depict the temperature profiles as shown in Fig. 4a with a temperature cut-off at 150 °C. With less available temperature information, the prediction of potential flashover occurrence will become more difficult. Therefore, in addition to being able to determine different building layout, the model also needs to have learning capabilities to correlate complex temperature information from other compartments to flashover conditions in the room of fire origin.

3 GENERIC FLASHOVER PREDICTION MODEL (FLASHNET)

3.1 Overall model structure

In this section, the overall structure of the generic flashover prediction model (FlashNet) is presented. As shown in Fig. 5, FlashNet is composed of two different modules. In the first module, the multivariate temperature inputs are transformed into graph representations and graph-structured data \( G = (X, W) \) are then formed where \( X \) is the node-attribute matrix and \( W \) is the adjacency matrix. For this study, the node-attribute matrix contains temperature information from the heat detectors. Consider Fig.1b as an example, there are six compartments. Given the current numerical setting with one heat detector in one compartment, \( X \in \mathbb{R}^{N \times D} \) where \( N \) is the number of compartments and \( D \) is the number of attributes. Since only temperature is used, \( D \) is equal to one and \( N \) is equal to six for the building structure as shown in Fig.1b. (rolling window is applied to extract temperature instances). For the adjacency matrix, \( W \in \mathbb{R}^{N \times N} \) in which \( w_{ij} \) is non-zero if compartment \( i \) and compartment \( j \) are connected. In principle, \( w_{ij} \) indicates the strength of the connection. If there is no connection between the two compartments, then \( w_{ij} \) is zero. Additional discussion is provided in the following subsections to explain how the adjacency matrix is determined.

![Diagram of FlashNet model](image)
In the second module, the graph data $G$ serves as inputs to the spatiotemporal graph convolution networks (STGCN) [29]. There are two STGCN blocks and they are used to capture the spatial and the temporal dependencies from the multivariate temperature data. Each STGCN block consists of two temporal gated convolution layers and one spatial graph convolution layer. For each temporal convolution layer, there are a 1-D convolution operation and a gated linear unit (GLU). Final feature representations are passed into the output layers to provide the prediction for the exact building structure and the potential occurrence of flashover with lead time of 30 s. It should be noted that the kernel size for both temporal and spatial convolutional operation is 3.

### 3.2 Adjacency matrix ($W$)

This section provides descriptions for the formulation of adjacency matrices and there are three separate steps. Firstly, a floor plan of a building structure is converted into a graph representation with nodes and edges. Figure 6a shows the graph representation of the six-compartment attached home (refer to Fig. 1b). As shown in Fig. 6a, each compartment is represented by a node and the corresponding opening or connection between two compartments is represented by an edge. The same procedure is carried out to generate the graph representations for the rest of the sixteen building structures.

In the second step, an adjacency matrix can be obtained for each of the building structures using the graph representations. It can be seen by visual inspection of the seventeen building structures from Appendix A that the formation of a building structure can be the combination of fifteen different compartments. Specifically, the combination can be a living room (L), kitchen (K), bedroom1 (BR1), bathroom1 (BA1), hall1 (H1), bathroom2 (BA2), bedroom2 (BR2), bathroom3 (BA3), den (Den), bedroom3 (BR3), family room (Fam), bedroom4 (BR4), hall2 (H2), hall3 (H3), or a dining room (D). Based on this combination, a square matrix with 15 elements can be formulated to describe the compartment relationship within a building structure. Figure 6b presents an adjacency matrix for the six-compartment attached home. The matrix elements, $w_{ij}$ for $i \neq j$, are determined based on the size of opening between two compartments (i.e., the height and the width of a door from a bedroom to the hallway). Since it is expected there will be an exchange of air from one compartment to another compartment in case of fire, the adjacency matrix is symmetric, indicating that temperature information is equally important. In principle, this adjacency matrix is denoted as an undirected graph [30]. For the diagonal elements, since the temperature information is crucial to correlate the potential occurrence of flashover, they are taken to be the maximum value of the non-diagonal elements. For the six-compartment building structure, the maximum value is 3. It should be noted that the formulation of the adjacency matrix is based on pre-defined information. However, additional research is on-going to establish a latent learning layer to allow the adjacency matrix to be determined based on the temperature information. In principle, this will help to eliminate the assumptions being made during the formulation of the adjacency matrix and to provide a more robust end-to-end prediction model.

![Figures 6. A schematic of a) a graph representation and b) an adjacency matrix for the six-compartment attached home (see Fig. 1b).](image-url)
After the adjacency matrices (AM) for the seventeen building structures are formulated, the last step involves the development of a geometric average adjacency matrix. The need of the geometric average adjacency matrix (GAAM) is due to the fact that the GNN algorithm cannot use multiple adjacency matrices for learning. To the best of the authors’ knowledge, the current study is the first attempt to overcome the difficulty of having multiple adjacency matrices. Figure 7 shows the GAAM for the selected seventeen structure and each of the matrix elements is determined based on the statistical mean of the seventeen AM. The GAAM is used together with the temperature instances (inputs) to train the generic prediction model.

### 3.3 Node attributes ($X$)

In this study, the data consists of temperature reading from the heat detectors and the objective of the model is to predict if there will be a potential occurrence of flashover from a particular compartment within a building structure in the future. Considering a building structure with $N$ heat detectors, the temperature data is denoted as $T = [T_0, T_5, \ldots, T_{FO}]$. The data sampling interval is 5 s and the notation, $FO$, is the time stamp for the transition to flashover. In order to allow fire fighters to have sufficient time to leave the compartment or find shelter, lead time, $x$, is considered and the available training data becomes $T = [T_0, T_5, \ldots, T_{FO-x}]$. Since it will take 10 s for fire fighters to travel about 3 m due to movement limitations (i.e., crawling) and poor visibility, the lead time is taken to be 30 s.

Sliding window is applied to construct instances from raw temperature data. An instance from a fire event is denoted as $I_i = \{S_i^1, S_i^2, \ldots, S_i^N\}$ where $S_i^j = (T_{t_0}^j, \ldots, T_{t_w}^j)$ with $i$ as the first time step of the sliding window and $w$ to be the window size. In this study, the window size is taken to be 5 mins. To account for heat detector failure at 150 °C, a failure moment ($t_{b_i}^j$) for signal $S_i^j$ is identified. If time $t \geq t_{b_i}^j$, the temperature reading, $T_{t_i}^j$, is replaced by a value of 0 °C, representing a loss of temperature signal. A masking layer is applied during the training process to neglect the zero values. Extracting all the $I_i$ from all fire events, the instance set of $\{I_0^1, I_1^1, \ldots, I_i^1, \ldots, I_0^M, I_1^M, \ldots, I_i^M\}$ is obtained to be the node attributes ($X$) where $M$ is the number of fire cases and the exact number of instances being used to train the current model will be provided in the following subsection. It should be noted that padding is used to make the node attributes (i.e., a vector with 15 elements) consistent with the adjacency matrix.

### 3.4 Labelling

Each instance is labelled to form the data samples and there are two types of labels. The first type of label is for classification task to determine the specific building structure. In general, this is relatively easy and it is just referring to the corresponding building structure (i.e., building #1, up to building #17) with one-hot encoding to which instances belong. The second type of label is associated with the classification task for predicting the potential occurrence of flashover and the instance is either labelled as flashover or non-

---

6 A fire event is considered as a fire cases and there are 78000 fire cases for seventeen different building structures in this study.
flashover based on the upper gas layer temperature. In this study, the threshold for the onset of flashover is selected to be 600 °C. For prediction lead time of 30 s, there are six flashover instances before the flashover condition is met and they are $I_{FO-30-w}$, $I_{FO-25-w}$, ..., $I_{FO-5-w}$. Given the fact that only two-fifth of the fire cases have flashover and a major of the instances from a flashover fire case are non-flashover, the current dataset has a data imbalance between flashover and non-flashover instances. In order to maintain data balance, six non-flashover instances are identified and they are taken to be $I_{FO-60-w}$, $I_{FO-55-w}$, ..., $I_{FO-35-w}$. With that, there are in total 493548 instances from 78000 fire cases.

### 3.5 Training and Testing

The dataset is divided into three subsets: training, validation, and testing sets. Specifically, a set of 4112 fire events worth of data samples are assigned to both validation and testing sets. The remaining 32905 (493548/12 - 8224) fire events are given to the training set. In order to facilitate the training process, the dataset is divided into batches and the batch size is determined to be 50. Using the RMSProp with a learning rate of 1e-3, the neural network weights and biases from the above equations are updated accordingly. SoftMax cross-entropy is used to measure the loss.

### 4 RESULTS

Table 4 shows the results of FlashNet for the prediction of potential flashover occurrence with lead time of 30 s and no lead time. Accuracy, precision, recall, and F1 score are used to evaluate the model performance. The mathematical expressions are the following: accuracy is (TP+TN)/(TP+FP+FN+TN) where TP, TN, FP, and FN are true positive, true negative, false positive, and false negative, respectively. Precision is TP/(TP+FP). Recall is TP/(TP+FN). F1 score is the ratio of 2*(Recall*Precision) and (Recall+Precision). Basically, these metrics offer additional insights about false predictions, either positive or negative, and the effects of data imbalance. As shown in the table, the overall accuracy for the model predicting the potential occurrence of flashover is approximately 93.1 %. Given the fact that only limited temperature information (i.e., up to 150 °C and no prior knowledge about the exact building structure) is used as inputs for prediction, the results are very encouraging. Comparing precision and recall, it can be seen that there are more false positives than false negatives. Predictions with lead time of 10 s are also provided. In general, the overall model performance decreases slightly and this is because 1) less temperature information is available and 2) the temperature behavior associated with the flashover occurrence become very similar.

Table 4 also shows the accuracy for classification of building structures based on temperature signals. In general, the model can predict the corresponding building structures excellently.

<table>
<thead>
<tr>
<th>Lead time</th>
<th>Accuracy</th>
<th>Precision</th>
<th>Recall</th>
<th>F1</th>
<th>Building Structures</th>
</tr>
</thead>
<tbody>
<tr>
<td>30 s</td>
<td>93.1 %</td>
<td>92.2 %</td>
<td>94.0 %</td>
<td>93.1 %</td>
<td>99.7 %</td>
</tr>
<tr>
<td>10 s</td>
<td>88.6 %</td>
<td>88.4 %</td>
<td>89.0 %</td>
<td>88.7 %</td>
<td></td>
</tr>
</tbody>
</table>

### 5 CONCLUSIONS

This paper provides a collection of single-story residential building structures including detached and attached homes and apartments. The floor area of these structures ranges from approximately 65 m² to 275 m² with about three to fourteen different compartments. Synthetic temperature data from a total number of 78000 cases are generated for model development.

The model structure of the generic flashover prediction model (FlashNet) is presented. The concept of graph-structure data is introduced. Examples for transforming the building structures into graph representations are illustrated. The geometric average adjacency matrix is formulated to facilitate the use of graph neural network. Using graph convolution operations, the spatial relationships and the temporal
dependencies from the multivariate temperature data are effectively captured. Results show that even when the available temperature information is limited (up to 150 °C), FlashNet can predict the potential occurrence of flashover in advance without the need of prior knowledge about the structure building. The overall accuracy is about 93 % with a possible forecast lead time of 30 s. This study contributes to the development of a generic ML-based model framework for a real-life flashover prediction model to enable data-driven fire fighting.

REFERENCES


It should be noted that the red lines are for visualization to identify the hallways in different building structures.