

**VIETNAM NATIONAL UNIVERSITY  
UNIVERSITY OF ENGINEERING AND TECHNOLOGY**



**Magneto @ DS&KTLab**

**ASPECT-ORIENTED SENTIMENT ANALYSIS  
FOR VIETNAMESE E-COMMERCE REVIEWS**

**STUDENT SCIENTIFIC RESEARCH REPORT**

**Faculty of Information Technology**

**HANOI - 2021**

**VIETNAM NATIONAL UNIVERSITY  
UNIVERSITY OF ENGINEERING AND TECHNOLOGY**

**Le Thi Phuong  
Le Minh Binh  
Bui Khanh Huyen  
Tran Khanh Hung**

**ASPECT-ORIENTED SENTIMENT ANALYSIS  
FOR VIETNAMESE E-COMMERCE REVIEWS**

**STUDENT SCIENTIFIC RESEARCH REPORT**

**Faculty of Information Technology**

**Supervisors:**

**MSc. Le Hoang Quynh**

**MSc. Can Duy Cat**

**HANOI - 2021**

# Abstract

Chúng tôi giới thiệu Graph Attention Network (GAT), mô hình mạng nơ-ron mới hoạt động trên dữ liệu dạng đồ thị, sử dụng các lớp tự quan tâm đã đặt mặt nạ để cải thiện những nhược điểm của các phương pháp trước đó dựa trên hỗn hợp đồ thị hoặc các ước lượng của chúng. Bằng cách xếp lớp trong đó các nút có thể quan tâm đến các đặc trưng của lân cận, chúng ta cho phép (ẩn dụ) chỉ định các trọng số khác nhau cho các nút trong một lân cận, mà không cần bất kỳ hoạt động ma trận chi phí (như đảo ngược) hoặc phụ thuộc vào việc biết cấu trúc đồ thị trước. Theo cách này, chúng tôi giải quyết một số vấn đề chính của các mạng nơ-ron đồ thị dựa trên tần số cùng lúc, và làm cho mô hình của chúng tôi dễ dàng áp dụng cho cả vấn đề dẫn chứng và dẫn điểm. Mô hình GAT của chúng tôi đã cho kết quả ??? trên bốn mô hình kiểm thử transductive và inductive: các tập dữ liệu Cora, Citeseer và Pubmed, cùng với tập dữ liệu tương tác protein-protein (ở đây, test graph bị ẩn trong quá trình học máy).

**Keywords:** *Graph Attention Networks(GAT), transductive, Vietnamese multi-aspect dataset.*

# Acknowledgements

Lời đầu tiên, chúng em xin cảm ơn Cô Nguyễn Thị Cẩm Vân và các thầy cô trong lab Data Science and Knowledge Technology Laboratory at University of Engineering and Technology. Các thầy cô đã luôn giúp đỡ và hướng dẫn chúng em viết ra bài báo này.

Đồng thời em cũng xin cảm ơn khoa Công nghệ thông tin - Trường Đại học Công nghệ đã tạo điều kiện để cho chúng em có được trải nghiệm thực tế trong môi trường nghiên cứu khoa học và tiếp thu nhiều kiến thức bổ ích.

# Declaration

Chúng tôi tuyên bố rằng tác phẩm của chúng tôi là do chính chúng tôi sáng tác và tác phẩm đó không được nộp cho bất kỳ bằng cấp hoặc trình độ chuyên môn nào khác. Chúng tôi xác nhận rằng tác phẩm được gửi là của riêng bạn, trừ khi tác phẩm đã tạo thành một phần của các ấn phẩm đồng tác giả đã được đưa vào. Đóng góp của chúng tôi và của các tác giả khác cho tác phẩm này đã được chỉ ra rõ ràng dưới đây. Chúng tôi xác nhận rằng tín dụng thích hợp đã được đưa ra trong bài viết này khi tài liệu tham khảo đã được thực hiện cho công việc của người khác.

Chúng tôi xác nhận rằng, theo hiểu biết tốt nhất của chúng tôi, bài viết của chúng tôi không vi phạm bản quyền của bất kỳ ai cũng như không vi phạm bất kỳ quyền sở hữu nào và rằng mọi ý tưởng, kỹ thuật, trích dẫn hoặc bất kỳ tài liệu nào khác từ tác phẩm của người khác được đưa vào luận án của tôi, đã xuất bản hoặc mặt khác, được thừa nhận đầy đủ theo các thông lệ tham chiếu tiêu chuẩn. Hơn nữa, trong phạm vi mà chúng tôi đã bao gồm tài liệu có bản quyền, chúng tôi xác nhận rằng chúng tôi đã nhận được sự cho phép bằng văn bản từ (những) chủ sở hữu bản quyền để đưa (những) tài liệu đó vào tác phẩm của chúng tôi và có đầy đủ quyền tác giả để cải thiện những tài liệu này.

Nhóm sinh viên

**Đoàn Văn Nguyên**  
**Nguyễn Đức Trọng**  
**Nguyễn Trần Đạt**

# Table of Contents

<b>Abstract</b>	<b>iii</b>
<b>Acknowledgements</b>	<b>iv</b>
<b>Declaration</b>	<b>v</b>
<b>Table of Contents</b>	<b>vi</b>
<b>1 Giới thiệu</b>	<b>1</b>
<b>2 MÔ HÌNH GRAPH ATTENTION NETWORK</b>	<b>4</b>
2.1 Đồ Thị Phân Lớp Tập Trung (Graph Attentional Layer)	4
2.2 So sánh với các nghiên cứu trước	7
<b>References</b>	<b>10</b>

# Chapter 1

## Giới thiệu

Mạng Nơ-ron Tích Chập (Convolutional Neural Networks - CNN) hiện đang được áp dụng để giải các bài toán phân loại hình ảnh (He et al., 2016), phân vùng ngữ nghĩa ảnh (Jegou et al., 2017) hoặc dịch máy (Gehring et al., 2016), trong đó dữ liệu nền được biểu diễn dưới cấu trúc giống mạng lưới. Các mô hình này tái sử dụng hiệu quả các bộ lọc cục bộ, với các tham số học, bằng cách áp dụng chúng cho tất cả các vị trí đầu vào. Tuy nhiên, nhiều tác vụ có liên quan đến dữ liệu không thể biểu diễn dưới dạng mạng lưới dưới dạng bất quy tắc. Bản dựng hình 3D, mạng xã hội, mạng viễn thông, mạng sinh học hoặc mạng kết nối não bộ là các ví dụ điển hình. Những loại dữ liệu này thường được biểu diễn dưới dạng đồ thị. Trong lĩnh vực công nghệ thông tin, đã có một số nghiên cứu về việc mở rộng mạng nơ-ron để xử lý các đồ thị. Các nghiên cứu ban đầu sử dụng mạng nơ-ron đệ quy để xử lý dữ liệu được biểu diễn dưới dạng đồ thị đường vòng được định hướng (Frasconi et al., 1998; Sperduti & Starita, 1997). Mạng Nơ-ron Đồ Thị (Graph Neural Network - GNN) được giới thiệu trong Gori et al. (2005) và Scarselli et al. (2009) là một mô hình chuẩn hóa của mạng nơ-ron đệ quy có thể trực tiếp xử lý các loại đồ thị bao quát hơn, ví dụ như đồ thị có chu trình, có hướng và không hướng. GNN bao gồm một quá trình lặp đi lặp lại, truyền trạng thái của nút cho đến khi đạt đối tượng; sau đó là một mạng nơ-ron, tạo ra một kết quả cho mỗi nút dựa trên trạng thái của nó. Ý tưởng này đã được Li et al. (2016) thừa hưởng và cải tiến bằng cách sử dụng đơn vị gated recurrent (Cho et al., 2014) trong bước truyền.

Trong khi đó, có một sự quan tâm tăng lên về việc mở rộng bộ lọc tổng quát cho không gian đồ thị. Những tiến bộ trong hướng này thường được phân loại thành các phương pháp phần tử và phương pháp không phần tử. Trên một bên, các phương pháp phần tử làm việc với một biểu diễn phần tử của đồ thị và đã được áp dụng thành công

trong mục tiêu phân loại nút. Trong Bruna et al. (2014), hoạt động tích chập được xác định trong tầng Fourier bằng cách tính phân tích độ riêng của ma trận Laplacian của đồ thị, dẫn đến việc tính toán có thể nặng và bộ lọc không phần tử tọa độ. Những vấn đề này đã được giải quyết bởi các công trình sau. Henaff et al. (2015) đã giới thiệu một tham số hóa của bộ lọc phần tử với các hệ số mượt để làm cho chúng có tọa độ phần tử. Sau đó, Defferrard et al. (2016) đề xuất sử dụng phép mở rộng Chebyshev của ma trận Laplacian để ước lượng bộ lọc, loại bỏ việc tính toán các vector riêng của Laplacian và cho ra các bộ lọc có tính địa lý, Kipf & Welling (2017) đã đơn giản hoá phương pháp trước bằng cách hạn chế bộ lọc hoạt động trong một vùng lân cận xung quanh mỗi nút. Tuy nhiên, trong tất cả các phương pháp đồng bộ hoá tần số trên, các bộ lọc được học phụ thuộc vào cơ sở vector riêng của Laplacian, điều này phụ thuộc vào cấu trúc đồ thị. Do đó, một mô hình được huấn luyện trên một cấu trúc cụ thể không thể được áp dụng trực tiếp cho một đồ thị với cấu trúc khác.

Trên một bên, các phương pháp không phổ hoạt (Duvenaud et al., 2015; Atwood & Towsley, 2016; Hamilton et al., 2017) xác định việc tích chập trực tiếp trên đồ thị, hoạt động trên nhóm các đỉnh gần về vị trí. Một trong những thách thức của các phương pháp này là xác định một toán tử hoạt động với các lân cận có kích thước khác nhau và giữ thuộc tính chia sẻ trọng lượng của CNN. Trong một số trường hợp, điều này yêu cầu học một ma trận trọng lượng cụ thể cho mỗi bậc đỉnh (Duvenaud et al., 2015), sử dụng các lũy thừa của ma trận chuyển đổi để xác định lân cận và học trọng lượng cho mỗi kênh đầu vào và bậc lân cận (Atwood & Towsley, 2016), hoặc trích xuất và chuẩn hóa các lân cận chứa số lượng cố định của đỉnh (Niepert et al., 2016). Monti et al. (2016) đã trình bày mô hình tổ hợp CNN (MoNet), một phương pháp vị trí cho phép tổng quát hóa kiến trúc CNN sang đồ thị. Gần đây, Hamilton et al. (2017) đã giới thiệu GraphSAGE, Một phương pháp tính biểu diễn của nút theo cách tuân theo. Kỹ thuật này hoạt động bằng cách mẫu một khu vực cố định kích thước của mỗi nút, sau đó thực hiện một trọng tải chung cho nó (như là trung bình của tất cả các vectơ đặc trưng của hàng xóm được mẫu, hoặc kết quả của việc cho chúng qua mạng nơ-ron tần suất). Phương pháp này đã cho kết quả tuyệt vời trên nhiều tiêu chuẩn thu hồi lớn.

Cơ chế tập trung đã trở thành gần như một tiêu chuẩn trong nhiều tác vụ dựa trên chuỗi (Bahdanau et al., 2015; Gehring et al., 2016). Một trong những lợi ích của các mạng tập trung là cho phép xử lý với kích thước đầu vào biến đổi, tập trung vào các phần quan trọng nhất của đầu vào để thực hiện quyết định. Khi sử dụng mạng tập trung để tính toán một biểu diễn của một chuỗi, nó thường được gọi là tập trung tự hoặc tập trung trong. Cùng với mạng nơ-ron lặp lại (RNNs) hoặc tích chập, tập trung tự đã chứng minh



là hữu ích cho các tác vụ như đọc máy (Cheng et al., 2016) và học biểu diễn câu (Lin et al., 2017). Tuy nhiên, Vaswani et al. (2017) cho thấy rằng tập trung tự không chỉ có thể cải thiện một phương pháp dựa trên RNNs hoặc tích chập, mà còn đủ để xây dựng một mô hình mạnh mẽ nhận được hiệu năng tiên tiến nhất trong tác vụ dịch máy.

Lấy cảm hứng từ các công trình mới nhất, chúng tôi giới thiệu một kiến trúc dựa trên chú ý để thực hiện phân loại nút của dữ liệu cấu trúc đồ thị. Ý tưởng là tính toán biểu diễn ẩn của mỗi nút trong đồ thị, bằng cách chú ý trên các nút lân cận của nó, theo một chiến lược chú ý tự. Kiến trúc chú ý có một số thuộc tính quan trọng: (1) hoạt động hiệu quả, vì nó có thể được parallel hoá giữa các cặp nút lân cận; (2) nó có thể được áp dụng cho các nút đồ thị có số bậc khác nhau bằng cách chỉ định các trọng lượng tùy ý cho các nút lân cận; và (3) mô hình có thể trực tiếp được áp dụng cho các vấn đề học theo một cách dựa trên thực tế, Bao gồm các nhiệm vụ mà mô hình phải tổng hợp đến các đồ thị hoàn toàn chưa từng thấy. Chúng tôi xác nhận giải pháp được đề xuất trên bốn mốc thử nghiệm khó: mạng trích dẫn Cora, Citeseer và Pubmed cũng như tập dữ liệu tương tác protein-protein theo phương pháp chỉ dẫn, đạt được hoặc phù hợp với kết quả tiên tiến nhất trong lĩnh vực để nổi bật tiềm năng của các mô hình dựa trên chú ý khi đối mặt với các đồ thị tùy ý.

Nên chú ý rằng, giống như Kipf & Welling (2017) và Atwood & Towsley (2016), công trình của chúng tôi cũng có thể được tái cấu trúc như một trường hợp cụ thể của MoNet (Monti et al., 2016). Hơn nữa, cách tiếp cận chúng tôi chia sẻ tính toán mạng nơ-ron qua các cạnh giống như cách tạo dạng của các mạng quan hệ (Santoro et al., 2017) và VAIN (Hoshen, 2017), trong đó các mối quan hệ giữa các đối tượng hoặc đại diện được tổng hợp theo cặp, bằng cách sử dụng một mục đích chung. Tương tự, mô hình chú ý được đề xuất của chúng tôi có thể được kết nối với các công trình của Duan et al. (2017) và Denil et al. (2017), sử dụng một hoạt động chú ý lân cận để tính toán các hệ số chú ý giữa các đối tượng khác nhau trong một môi trường. Các phương pháp liên quan khác bao gồm nhúng tính linearly cục bộ (LLE) (Roweis & Saul, 2000) và các mạng nhớ (Weston et al., 2014). LLE chọn một số cố định của các láng giềng xung quanh mỗi điểm dữ liệu, và học một trọng số trọng tâm cho mỗi điểm lân cận để tái tạo mỗi điểm như là tổng trọng tâm của các điểm lân cận của nó. Bước tối ưu hóa thứ hai trích xuất nhúng đặc trưng của điểm. Mạng nhớ cũng chia sẻ một số kết nối với công việc của chúng tôi, đặc biệt là nếu chúng tôi hiểu khu vực xung quanh một nút như là bộ nhớ, được sử dụng để tính toán các đặc trưng của nút bằng cách chú ý về giá trị của nó, và sau đó được cập nhật bằng cách lưu trữ các đặc trưng mới tại cùng vị trí

## Chapter 2

# MÔ HÌNH GRAPH ATTENTION NETWORK

Trong phần này, chúng tôi sẽ trình bày các kiến thức nền tảng được sử dụng để xây dựng các mạng quan sát đồ thị tùy chọn (bằng cách gộp lớp này) và liệt kê các hiệu quả về mặt lý thuyết và thực tế cùng các hạn chế so với công trình liên quan đến lĩnh vực xử lý đồ thị nơ-ron trước đây."

### 2.1 Đồ Thị Phân Lớp Tập Trung (Graph Attentional Layer)

Chúng ta bắt đầu bằng viết một phân lớp đồ thị tập trung đơn lẻ, đây là lớp duy nhất được sử dụng trong toàn bộ các mô hình GAT trong các thử nghiệm của chúng tôi. Thiết kế chú ý cụ thể được chúng tôi sử dụng gần gũi với công trình của Bahdanau et al. (2015) - nhưng khung làm việc là agnostic với lựa chọn cụ thể của mức độ chú ý.

Đầu vào của lớp của chúng tôi là một tập hợp các đặc điểm nút,  $\mathbf{h} = \{\vec{h}_1, \vec{h}_2, \dots, \vec{h}_N\}$ ,  $\vec{h}_i \in \mathbb{R}^F$ , trong đó  $N$  là số lượng nút và  $F$  là số lượng đặc điểm trong mỗi nút. Lớp tạo ra một tập hợp các đặc điểm nút mới (có thể khác với số lượng  $F$ ),  $\mathbf{h}' = \{\vec{h}'_1, \vec{h}'_2, \dots, \vec{h}'_N\}$ ,  $\vec{h}'_i \in \mathbb{R}^F$ , như đầu ra của nó.

Để có đủ năng lực biểu diễn để chuyển đổi đặc trưng đầu vào thành đặc trưng cấp cao hơn, ít nhất một biến đổi tuyến tính có thể học là cần thiết. Để đạt được mục đích đó, làm một bước đầu tiên, một biến đổi tuyến tính chung, được tham số hóa bởi *ma trận trọng số*,  $\mathbf{W} \in \mathbb{R}^{F' \times F}$ , được áp dụng cho mỗi nút. Chúng ta sau đó thực hiện tự

quan tâm trên các nút - một mục đích quan tâm chung  $a: \mathbb{R}^{F'} \times \mathbb{R}^F \rightarrow \mathbb{R}$  tính hệ số chú ý

$$e_{ij} = a(\mathbf{W}\vec{h}_i, \mathbf{W}\vec{h}_j)$$

cho thấy sự quan trọng của đặc trưng của nút  $j$  đối với nút  $i$ . Trong biểu diễn tổng quát nhất, mô hình cho phép mỗi nút chú ý tới mọi nút khác, *loại bỏ tất cả thông tin cấu trúc*. Chúng ta tiếp cận cấu trúc đồ thị qua việc thực hiện sự chú ý được che đậy - Chỉ tính  $e_{ij}$  cho các nút  $j \in N_i$ , trong đó  $N_i$  là một khu vực xung quanh nút  $i$  trong đồ thị. Trong tất cả các thí nghiệm của chúng tôi, đó sẽ chính là các nút đồng bậc đầu tiên của  $i$  (bao gồm  $i$ ). Để dễ dàng so sánh các hệ số giữa các nút khác nhau, chúng tôi chuẩn hóa chúng theo tất cả các lựa chọn của  $j$  bằng hàm softmax:

$$\alpha_{ij} = \text{softmax}_j(e_{ij}) = \frac{\exp(e_{ij})}{\sum_{k \in N_i} \exp(e_{ik})}$$

Trong các thí nghiệm của chúng tôi, cơ chế chú ý  $a$  là một mạng nơ-ron feedforward đơn tầng, được tham số hóa bởi một vectơ trọng số  $\vec{a} \in \mathbb{R}_{2F'}$  và áp dụng phi tuyến LeakyReLU (với đầu vào đối mạng âm  $\alpha = 0, 2$ ). Để mở rộng ra, các hệ số được tính bởi cơ chế chú ý (minh họa bởi Hình 1 (trái)) có thể được biểu diễn như sau:

$$\alpha_{ij} = \frac{\exp(\text{LeakyReLU}(\vec{a}^T [\mathbf{W}\vec{h}_i \| \mathbf{W}\vec{h}_j]))}{\sum_{k \in N_i} \exp(\text{LeakyReLU}(\vec{a}^T [\mathbf{W}\vec{h}_i \| \mathbf{W}\vec{h}_k]))}$$

$\cdot^T$  đại diện cho phép chuyển vị, và  $\|$  là thao tác nối tiếp.

Sau khi được tính toán, các hệ số quan tâm được chuẩn hóa được sử dụng để tính tổng trọng số của các đặc trưng tương ứng với chúng, để dùng làm đặc trưng đầu ra cuối cùng cho mỗi nút (sau có thể là

Áp dụng một hàm phi tuyến,  $\sigma$ ):

$$\vec{h}_i' = \sigma\left(\sum_{j \in N_i} \alpha_{ij} \mathbf{W}\vec{h}_j\right)$$

Để stabilizr quá trình học tự quan tâm, chúng tôi đã tìm thấy mở rộng cơ chế của chúng tôi để sử dụng quan tâm đa đầu có lợi, tương tự như Vaswani et al. (2017). Cụ thể, K cơ chế quan tâm độc lập thực hiện biến đổi của Phương trình 4, và sau đó các đặc trưng của chúng được nối tiếp, dẫn đến biểu diễn đặc trưng đầu ra sau:

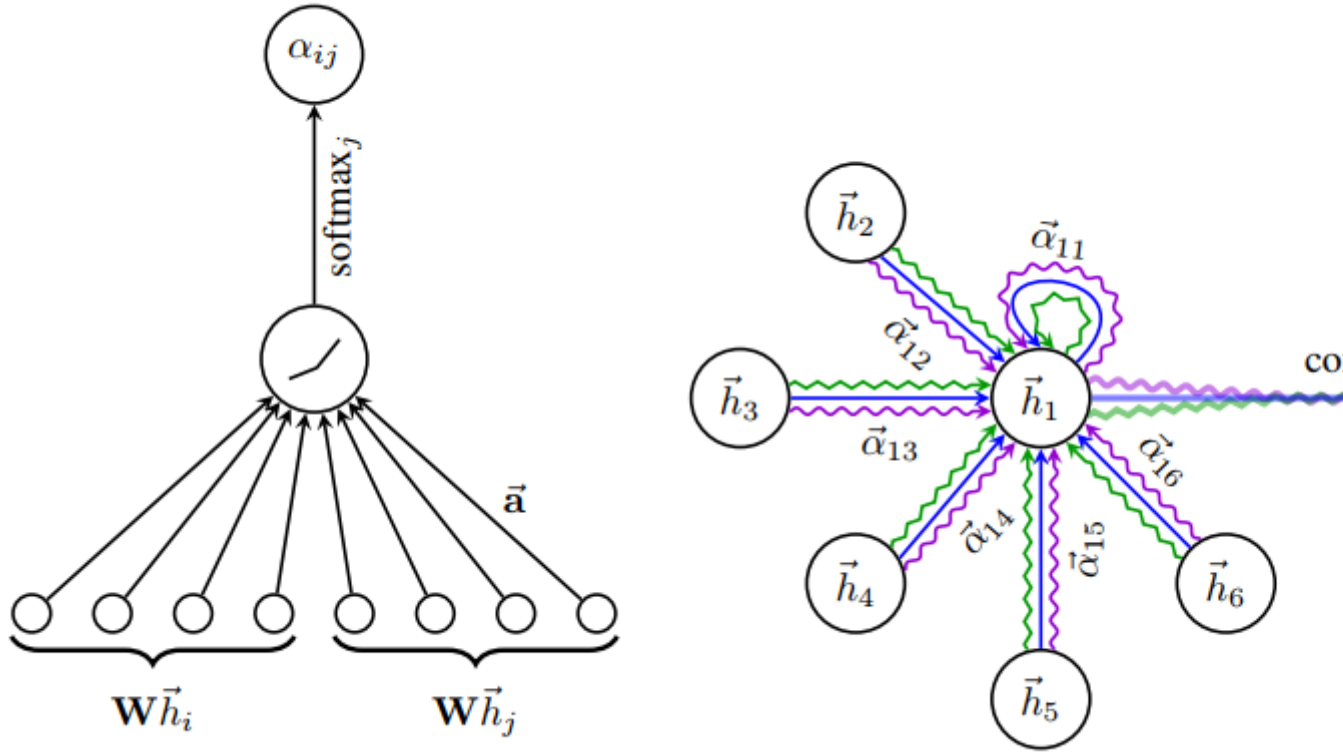


Figure 2.1: **Left:** Cơ chế quan tâm  $a(\mathbf{W}\vec{h}_i\|\mathbf{W}\vec{h}_j)$  được sử dụng bởi mô hình của chúng tôi, được tham số hóa bởi một vectơ trọng số  $\vec{a} \in \mathbb{R}^{2F'}$ , áp dụng một hoạt động LeakyReLU. Bên phải: Một minh họa của quan tâm đa đầu (với  $K = 3$  đầu) bởi nút 1 trên khu vực xung quanh của nó. Các kiểu mũi tên và màu sắc khác nhau đại diện cho các tính toán quan tâm độc lập. Các đặc trưng được tổng hợp từ mỗi đầu được nối tiếp hoặc trung bình để đạt được  $\vec{h}'_1$ ,

$$\vec{h}'_i = \big\|_{k=1}^K \sigma \left( \sum_{j \in \mathcal{N}_i} \alpha_{ij}^k \mathbf{W}^k \vec{h}_j \right)$$

Trong đó,  $\|$  đại diện cho việc nối tiếp,  $\alpha_{ij}^k$  là các hệ số quan tâm đã chuẩn hóa được tính bởi cơ chế quan tâm  $k$ -th ( $a^k$ ), và  $\mathbf{W}^k$  là ma trận trọng số biến đổi tuyến tính đầu vào tương ứng. Lưu ý rằng, trong tình huống này, kết quả trả về cuối cùng,  $\mathbf{h}'$ , sẽ bao gồm  $K F'$  đặc trưng (thay vì  $F'$ ) cho mỗi nút.

Đặc biệt, nếu chúng ta thực hiện quan tâm đa đầu trên lớp cuối cùng (dự đoán) của mạng, việc nối tiếp không còn hợp lý - thay vào đó, chúng tôi sử dụng trung bình, và hoãn áp dụng nonlinearity cuối cùng (thường là một softmax hoặc sigmoid logistic cho các vấn đề phân loại) cho đến khi đó:

$$\vec{h}'_i = \sigma \left( \frac{1}{K} \sum_{k=1}^K \sum_{j \in \mathcal{N}_i} \alpha_{ij}^k \mathbf{W}^k \vec{h}_j \right)$$

Quá trình tập hợp của một lớp quan tâm đồ thị đa đầu được minh họa bởi Hình 1 (bên phải).

## 2.2 So sánh với các nghiên cứu trước

Lớp quan sát đồ thị mô tả trong phần 2.1 trực tiếp giải quyết một số vấn đề xuất hiện trong các phương pháp trước đó về mô hình hóa dữ liệu cấu trúc đồ thị:

- Tính toán, nó rất hiệu quả: hoạt động của lớp tự quan tâm có thể được đồng bộ hóa trên tất cả các cạnh và tính toán của đặc trưng đầu ra có thể được đồng bộ hóa trên tất cả các nút. Không cần phải tính toán các phép biến đổi ma trận giống như tính toán *eigendecompositions* hoặc các phép toán ma trận tương tự chỉ phí cao. Độ phức tạp của một đầu quan tâm GAT tính toán đặc trưng  $F'$  có thể biểu diễn dưới dạng  $O(|V|FF' + |E|F')$ ,  $F$  là số lượng các đặc trưng đầu vào,  $|V|$  và  $|E|$  là số lượng nút và cạnh trong đồ thị, tương ứng. Độ phức tạp này tương đương với các phương pháp cơ sở như Mạng Convolutional đồ thị (GCN) (Kipf & Welling, 2017). Áp dụng nhiều đầu chú ý tăng yêu cầu lưu trữ và tham số lên một yếu tố  $K$ , trong khi tính toán của các đầu riêng biệt là hoàn toàn độc lập và có thể song song hóa.
- Khác với GCNs, mô hình của chúng tôi cho phép (ẩn định) gán sự quan trọng khác

nhau cho các nút trong một khu vực giống nhau, cho phép mô hình có khả năng lớn hơn. Ngoài ra, phân tích trọng số chú ý đã học có thể dẫn đến lợi ích trong khả năng hiểu biết, như trong lĩnh vực dịch máy (ví dụ, phân tích chất lượng của Bahdanau et al. (2015))

- Mechanism attention được áp dụng theo cách chung cho tất cả các cạnh trong đồ thị và do đó không phụ thuộc vào truy cập trước tiên đến cấu trúc đồ thị toàn cầu hoặc (các đặc trưng của) tất cả nút (một hạn chế của nhiều kỹ thuật trước đó). Điều này có nhiều hậu quả mong muốn:
  - Không yêu cầu đồ thị phải là không hướng (chúng ta có thể để trống tính toán  $ij$  nếu cạnh  $j \rightarrow i$  không tồn tại).
  - Cách tiếp cận của chúng ta cho phép áp dụng trực tiếp cho huấn luyện tại chỗ
    - bao gồm các nhiệm vụ nơi mô hình được đánh giá trên các đồ thị hoàn toàn chưa từng nhìn thấy trong quá trình huấn luyện.
- Kỹ thuật chú ý được áp dụng theo cách chung đến tất cả các cạnh trong đồ thị và do đó, nó không phụ thuộc vào truy cập trước tới cấu trúc đồ thị toàn cầu hoặc tính năng của tất cả nút (giới hạn của nhiều kỹ thuật trước đó). Điều này có nhiều kết quả mong muốn: cách tiếp cận mới của Hamilton et al. (2017) lấy mẫu một vùng lưới cố định của mỗi nút để giữ cho chân dung tính toán giống nhau, điều này không cho phép truy cập tới toàn bộ vùng lưới trong quá trình suy diễn. Ngoài ra, kỹ thuật này đạt được một số kết quả mạnh mẽ nhất khi sử dụng tập trung vùng lưới dựa trên LSTM (Hochreiter & Schmidhuber, 1997). Điều này giả định tồn tại một thứ tự nút liên tục qua các vùng lưới và tác giả đã sửa chữa nó bằng cách truyền vào các dãy ngẫu nhiên được sắp xếp cho LSTM. Cách thức của chúng tôi không gặp phải bất kỳ vấn đề nào của những vấn đề đó - nó hoạt động với toàn bộ khu vực lân cận (với chi phí tính toán biến đổi, vẫn cùng mức với các phương pháp như GCN), và không giả định bất kỳ sắp xếp nào trong nó.
- GAT có thể được tái cấu trúc thành một trường hợp cụ thể của MoNet (Monti et al., 2016). Cụ thể hơn thiết lập hàm tọa độ giả là  $u(x, y) = f(x) \| f(y)$ , trong đó  $f(x)$  là đặc trưng (có thể được chuyển đổi bằng MLP) của nút  $x$  và  $\|$  là sự nối tiếp; và hàm trọng lượng là  $\omega_j(u) = \text{softmax}(MLP(u))$  (với softmax được thực hiện trên toàn bộ lân cận của một nút), sẽ kiến tạo tử patch của MoNet tương tự với chúng tôi. Tuy nhiên, cần lưu ý rằng, so với các phiên bản MoNet đã được xem xét trước đây, mô hình của chúng tôi sử dụng các tính năng nút để tính toán độ tương tự, thay vì các thuộc tính cấu trúc của nút (giả sử biết trước cấu trúc biểu đồ).

Chúng tôi đã có thể tạo ra một phiên bản của lớp GAT sử dụng các hoạt động ma trận rỗng, giảm độ phức tạp lưu trữ thành tuyến tính trong số lượng nút và cạnh và cho phép thực hiện các mô hình GAT trên các tập dữ liệu đồ thị lớn hơn. Tuy nhiên, khung quản lý xử lý tensor mà chúng tôi sử dụng chỉ hỗ trợ nhân ma trận rỗng cho tensor rank-2, điều này hạn chế khả năng nhóm của lớp như hiện tại đã triển khai (đặc biệt là cho các tập dữ liệu có nhiều đồ thị). Định hướng giải quyết hạn chế này là quan trọng đối với công việc trong tương lai. Tùy thuộc vào tính chính xác của cấp trúc đồ thị, GPU có thể không cung cấp lợi ích hiệu suất lớn so với CPU trong những trường hợp rỗng. Nên chú ý rằng kích thước của “Khung nhìn” của mô hình của chúng tôi được giới hạn bởi độ sâu của mạng (tương tự cho GCN và các mô hình tương tự). Kỹ thuật như kết nối bỏ qua (He et al., 2016) có thể được áp dụng dễ dàng để mở rộng độ sâu, tuy nhiên. Cuối cùng việc lập trình song song cho tất cả các cạnh đồ thị, đặc biệt là theo cách phân tán, có thể gặp nhiều tính toán trùng lặp, vì khu vực lân cận thường rất tương đồng trong các đồ thị quan tâm.

# References