Combining Graph Neural Network with Attention Mechanism in Sequence-Based Music Recommendation System

1st Rama Aulia Gemilang
School of Computing
Universitas Telkom
Bandung, Indonesia
gemilangrama@student.telkomuniversity.ac.id

2nd Agung Toto Wibowo
School of Computing
Universitas Telkom
Bandung, Indonesia
agungtoto@telkomuniversity.ac.id

***Abstract*—Music recommendations have become an important part of everyday life. With the increasing amount of music available, it has become increasingly difficult to find suitable music content. To address this issue, recommendation systems have emerged as a solution to personalize song suggestions for users. Traditional recommendation systems often fail to capture sequential patterns of user interaction. To understand these patterns, recommendation systems are based on a sequential approach. However, traditional sequential recommendation system models do not sufficiently capture dependencies on long-term temporal patterns. In this study, we apply a combination of graph neural networks with attention mechanisms to the Music4All dataset to address sequential music recommendations. This approach, which adapts the GASM framework, is applied for the first time to this music dataset, which has greater diversity and user distribution, providing new insights into the effectiveness of this model in the music recommendation domain. Graph neural networks capture the relationships between users and items, as well as between items themselves, while the attention mechanism captures user interest patterns over time within a session. We also incorporate long-term interests and dynamic interests to enrich the context of user preferences. By applying this model to the Music4All dataset, the model can recommend the next music track with a hit ratio evaluation ranging from 32.69 to 49.16 and a mean reciprocal rank ranging from 23.34 to 30.71. These results show that our approach can effectively capture sequential patterns in music in large-scale datasets.**

***Index Terms*—Sequential recommendation system, music, graph neural network, attention mechanism, preference.**

I. Introduction

In this digital age, music has become almost inseparable from everyday life. This is followed by the growing music production. The development of music production makes the number of songs available to consumers also become more and more [1]. This makes it more difficult for consumers to find music content that suits their preferences because they will spend a lot of time searching for music that suits their tastes, which can reduce the user experience in finding music [2]. To reduce the burden on consumers in finding music content, recommendation systems play an important role in this regard.

A recommendation system is a system that can personalize with the aim of providing product suggestions that may be of interest to users [3]. The main goal is to improve the user experience. Improved user experience in recommendation systems can be achieved from user interactions over time. In this case the recommendation can be based on a time sequence approach called a sequential recommendation system.

Sequential recommendation systems are tasked with analyzing user interaction history to predict consumer preferences [4]. User interactions can be a history of accessing sites, listening to songs, watching videos, etc. The interaction history will usually form a time sequence pattern following the transition from one item to the next. Using interaction time sequence patterns, the recommendation system will accurately understand user preferences. One of the implementations of sequential recommendation systems is through traditional approaches, such as markov models that provide item predictions by considering the last interaction. However, such models cannot handle long-term sequence dependency [5]. In addition, there are latent factor models that work by learning latent representations to analyze transitions between items. However, transition models often vary across different data, which makes them in a position to capture complex transition relationships [5].

To overcome sequence dependency, researchers have conducted several experiments, one of which uses a neural network graph representing user interactions with items and attention mechanisms that provide important weights in the order of user history. Research by Chang et al. [6], has integrated a graph neural network model as a graph of user relationships with items and attention mechanisms as weights in assessing preferences in short-term transitions. The integration of the model showed better results than the baseline in handling the dependence of time sequence [6].

In this study, we adapt the Graph-based Attentive Sequential model with Metadata (GASM) approach [7]. The model works by combining a graphical neural network with an attention mechanism, where the graphical neural network forms connections between users and items, as well as between items and items, while the attention mechanism forms short-term preferences by weighting items based on changes that occur over time during a session. In addition, other preferences are built dynamically over the long term, creating complex