A compass towards equity: a data analysis framework to capture children’s behaviour in the playground context
Nasri, M.

A Compass Towards Equity

A Data Analysis Framework to Capture Children's Behaviour in the Playground Context

Maedeh Nasri

A Compass Towards Equity

Proefschrift

ter verkrijging van
de graad van doctor aan de Universiteit Leiden,
op gezag van rector magnificus prof.dr.ir. H. Bijl,
volgens besluit van het college voor promoties
te verdedigen op dinsdag 3 december 2024
klokke 14:30 uur

door

Maedeh Nasri
geboren te Ramsar, Iran
in 1994.

Promotor:
Prof.dr. C. Rieffe
Co-promotores:
Dr. A. Koutamanis Delft University of Technology
Dr. M. Baratchi
Promotiecommissie:
Prof.dr. H.E. Hulst (Wetenschappelijk directeur psychologie/voorzitter)
Prof.dr. S. Giest
Prof.dr.ir. W. Kraaij
Dr.ir. D. Reidsma University of Twente
Dr. G. Veiga University of Evora, Portugal

This dissertation represents independent research funded by the Dutch Research Council (NWO, grant number: AUT.17.007) and Leiden-Delft-Erasmus Centre for BOLD Cities (grant number: BC2019-1).

This dissertation has been approved by:

Promotores:	Prof.dr. C. Rieffe
	Dr. A. Koutamanis
	Dr. M. Baratchi

This dissertation is dedicated to *my life partner, Ashkan and my family*

Summary

Playgrounds help children develop essential physical and social skills through playing and socializing with their peers and supervisors. However, playgrounds may also include obstacles that hinder children's development, specifically for those with communication difficulties. Addressing these challenges requires a deep understanding of children's needs in playground settings.

By adopting wearable sensing technology, we aimed to design a data analysis framework and deepen our knowledge of children's behavior in playgrounds. Developing a practical playground analysis framework requires addressing three main characteristics of playgrounds: multiple interconnected environments, differences in individual experiences, and spatio-temporal dynamics. Through various research questions, we have addressed these features as follows.

- **Feature 1: Multiple Environments.**
  In playgrounds, children interact with three interconnected environments: physical environment, social environment, and cultural environment. To incorporate these multiple environments in our data analysis framework, we used modern sensing technologies, i.e., proximity tags, GPS loggers, and accelerometers, to capture children's behavior in schoolyards. This method further enabled us to identify three affordances, i.e., physical, social, and cultural, which a child interacts with according to its capacities and needs, i.e., effectivities. We further developed a spatio-temporal metric that measures the impact of the physical environment on children's social networks in the playground context.

- **Feature 2: Individual Experiences.**
  Children have different capacities and needs in their interactions with their environments. We specifically studied these differences in individual experiences by focusing on children with and without autism. Individual experiences

are included through self-report data and peer nomination reports alongside sensor measurements to obtain children's perspectives on their behavior in playgrounds. Our results show the importance of addressing differences in effectivities either at the group level or at the individual level in relation to the affordances of the environment.

- **Feature 3: Spatio-temporal Dynamics.**
  Children's activities constantly change in playgrounds during play. Analyzing these spatio-temporal dynamics enables a deeper understanding of children's behavior in their social environment. We specifically focused on capturing complex interactions beyond face-to-face contact by utilizing artificial intelligence models to model group behavior. Our proposed method identifies group behaviors using spatio-temporal data in constrained environments such as university campuses or schoolyards.

The present dissertation integrates cutting-edge technological advancements with multidisciplinary collaboration to design a framework for analyzing children's behavior in playgrounds. This framework offers stakeholders a valuable tool for analyzing individual and group-level challenges in micro-communities, such as schoolyards, nursing homes, and sports clubs.

Samenvatting

Speeltuinen helpen kinderen essentiële fysieke en sociale vaardigheden te ontwikkelen door te spelen en te socialiseren met hun leeftijdsgenoten en begeleiders. Speeltuinen kunnen echter ook obstakels bevatten die de ontwikkeling van kinderen belemmeren, met name voor kinderen met communicatieproblemen. Om deze uitdagingen aan te pakken, is een diepgaand begrip van de behoeften van kinderen in speeltuinomgevingen vereist.

Door draagbare sensortechnologie te gebruiken, wilden we een data-analysekader ontwerpen en onze kennis van het gedrag van kinderen op speeltuinen verdiepen. Om een praktisch analysekader voor speeltuinen te ontwikkelen, moeten we drie hoofdkenmerken van speeltuinen aanpakken: meerdere onderling verbonden omgevingen, verschillen in individuele ervaringen en ruimtelijk-temporele dynamiek. Via verschillende onderzoeksvragen hebben we deze kenmerken als volgt aangepakt.

- **Kenmerk 1: Meerdere omgevingen.**
  Op speeltuinen interacteren kinderen met drie onderling verbonden omgevingen: fysieke omgeving, sociale omgeving en culturele omgeving. Om deze meerdere omgevingen in ons data-analysekader op te nemen, hebben we moderne sensortechnologie gebruikt, d.w.z. nabijheidstags, GPS-loggers en accelerometers, om het gedrag van kinderen op schoolpleinen vast te leggen. Deze methode stelde ons verder in staat om drie affordances te identificeren, namelijk fysiek, sociaal en cultureel, waarmee een kind omgaat op basis van zijn capaciteiten en behoeften, oftewel effectiviteit. We ontwikkelden verder een spatiotemporele metriek die de impact van de fysieke omgeving op de sociale netwerken van kinderen in de context van de speeltuin meet.

- **Kenmerk 2: Individuele ervaringen.**
  Kinderen hebben verschillende capaciteiten en behoeften in hun interacties met hun omgeving. We bestudeerden specifiek deze verschillen in individuele

ervaringen door ons te richten op kinderen met en zonder autisme. Individuele ervaringen worden opgenomen via zelfrapportagegegevens en peernominatierapporten naast sensormetingen om de perspectieven van kinderen op hun gedrag op speeltuinen te verkrijgen. Onze resultaten tonen het belang aan van het aanpakken van verschillen in effectiviteit, hetzij op groepsniveau of op individueel niveau, in relatie tot de affordances van de omgeving.

- **Kenmerk 3: Spatiotemporele dynamiek.**
  De activiteiten van kinderen veranderen voortdurend op speeltuinen tijdens het spelen. Door deze spatiotemporele dynamiek te analyseren, krijgen we een dieper inzicht in het gedrag van kinderen in hun sociale omgeving. We hebben ons specifiek gericht op het vastleggen van complexe interacties die verder gaan dan face-to-face contact door gebruik te maken van modellen voor kunstmatige intelligentie om groepsgedrag te modelleren. Onze voorgestelde methode identificeert groepsgedrag met behulp van spatio-temporele gegevens in beperkte omgevingen zoals universiteitscampussen of schoolpleinen.

Het huidige proefschrift integreert geavanceerde technologische ontwikkelingen met multidisciplinaire samenwerking om een raamwerk te ontwerpen voor het analyseren van het gedrag van kinderen op speelplaatsen. Dit raamwerk biedt belanghebbenden een waardevolle tool voor het analyseren van uitdagingen op individueel en groepsniveau in microgemeenschappen, zoals schoolpleinen, verpleeghuizen en sportclubs.

List of Publications

This dissertation is based on the following publications:

Journal Papers

[J1] M. Nasri, Y.-T. Tsou, A. Koutamanis, M. Baratchi, S. Giest, D. Reidsma, and C. Rieffe, “A novel data-driven approach to examine children’s movements and social behaviour in schoolyard environments,” *Children*, vol. 9, no. 8, p. 1177, 2022

[J2] M. Nasri, M. Baratchi, Y.-T. Tsou, S. Giest, A. Koutamanis, and C. Rieffe, “A novel metric to measure spatio-temporal proximity: a case study analyzing children’s social network in schoolyards,” *Applied Network Science*, vol. 8, no. 1, p. 50, 2023

[J3] M. Nasri, Z. Fang, M. Baratchi, G. Englebienne, S. Wang, A. Koutamanis, and C. Rieffe, “A gnn-based architecture for group detection from spatio-temporal trajectory data,” in *International Symposium on Intelligent Data Analysis*, pp. 327–339, Springer, 2023

[J4] M. Nasri, T. Maliappis, C. Rieffe, and M. Baratchi, “T-dante: Detecting group behaviour in spatio-temporal trajectories using context information,” in *International Symposium on Intelligent Data Analysis*, pp. 28–39, Springer, 2024

[J5] Y.-T. Tsou, M. Nasri, B. Li, E. M. Blijd-Hoogewys, M. Baratchi, A. Koutamanis, and C. Rieffe, “Social connectedness and loneliness in school for autistic and allistic children,” *Autism*, p. 13623613241259932, 2024

[J6] *(under review)* M. Nasri, M. Baratchi, A. Koutamanis, C. Rieffe, “SiamCircle: Trajectory Representation Learning in Free Settings”, IDA, 2025.

Conference Paper

[C1] M. Nasri, S. Giest, Y.-T. Tsou, M. Baratchi, A. Koutamanis, and C. Rieffe, "Role of movement data in creating inclusive school settings for students," in *Data for Policy*, 2022

Other Publications

[O1] A. Eichengreen, M. van Rooijen, L.-M. van Klaveren, M. Nasri, Y.-T. Tsou, A. Koutamanis, M. Baratchi, and C. Rieffe, "The impact of loose-parts-play on schoolyard social participation of children with and without disabilities: A case study," *Child: care, health and development*, vol. 50, no. 1, p. e13144, 2024

[O2] A. Eichengreen, Y.-T. Tsou, M. Nasri, L.-M. van Klaveren, B. Li, A. Koutamanis, M. Baratchi, E. Blijd-Hoogewys, J. Kok, and C. Rieffe, "Social connectedness at the playground before and after covid-19 school closure," *Journal of applied developmental psychology*, vol. 87, p. 101562, 2023

[O3] R. Laanen, M. Nasri, R. van Dijk, M. Baratchi, A. Koutamanis, and C. Rieffe, "Automated classification of pre-defined movement patterns: A comparison between gnss and uwb technology," *arXiv preprint arXiv:2303.07107*, 2023

Grants

- "SmartPlay: A smartwatch software for measuring play with a data-driven, interdisciplinary approach" (2023; 1150 hours of Research Software Engineer support). Assigned by the Netherlands eScience Center to C. Rieffe, Y.T. Tsou, M. Nasri, E.M.A. Blijd-Hoogewys, A. Koutamanis, and B. Li.

Table of Contents

Summary **vii**

Samenvatting **ix**

List of Publications **xi**

1 General Introduction **1**

1.1 Background 2

1.1.1 Children in Playgrounds 2

1.2 Playground Analysis Framework 3

1.3 This Thesis 4

1.3.1 Organization of Thesis 8

2 A Novel Data-driven Approach to Examine Children's Movements and Social Behaviour in Schoolyard Environments **9**

2.1 Introduction 10

2.1.1 Affordances and Effectivities in Physical, Social, and Cultural Environments in Schoolyards 11

2.1.2 Present Study 14

2.2 Methodology 15

2.2.1 Selection 15

2.2.2 Reconnaissance 16

2.2.3 Participants 16

2.2.4 Validation of Measures Obtained through Sensor Techniques 18

2.2.5 Variables and Measures . . . 19
2.2.5.1 GPS Loggers . . . 19
2.2.5.2 Proximity Tags . . . 20
2.2.5.3 MMR sensors . . . 20
2.3 Results . . . 21
2.3.1 Physical Affordances . . . 21
2.3.2 Social Affordances . . . 22
2.3.3 Cultural Affordances . . . 26
2.4 Discussion & Conclusions . . . 27
2.4.1 Limitations and Future Directions . . . 30

3 A Novel Metric to Measure Spatio-temporal Proximity: A Case Study Analyzing Children's Social Network in Schoolyards 33
3.1 Introduction . . . 34
3.2 Related work . . . 38
3.2.1 Location-Based Social Network (LBSN) Analysis . . . 38
3.2.2 Spatio-temporal Pattern Mining . . . 39
3.3 Problem Definition . . . 40
3.4 Methodology . . . 40
3.4.1 Background and Preliminaries . . . 41
3.4.2 Spatio-temporal Proximity . . . 43
3.5 Experiments . . . 45
3.5.1 Experiment Settings . . . 46
3.5.2 Schoolyard Dataset . . . 48
3.5.2.1 Data Preparation and Pre-processing . . . 49
3.5.3 Evaluation Metrics . . . 50
3.5.4 Results and Discussion . . . 51
3.5.4.1 Clustering Performance Evaluation . . . 51
3.5.4.2 In-Depth analysis of Schoolyard Data . . . 52
3.6 Conclusions . . . 59

4 Social Connectedness and Loneliness in School for Autistic and Non-autistic Children 61
4.1 Introduction . . . 62
4.1.1 Social connectedness and loneliness in school for autistic and allistic children . . . 63
4.1.2 Social Connectedness of Autistic Pupils . . . 63
4.1.3 Loneliness in Autistic Pupils . . . 64
4.2 Present Study . . . 65

4.3 Methods 67
4.3.1 Participants 67
4.3.2 Measures and Procedures 68
4.3.2.1 Measures of Physical Connectedness via Wearable Proximity Sensors 69
4.3.2.2 Measures of Emotional Connectedness via Peer Nomination 72
4.3.2.3 Loneliness in School via Self-Reports 73
4.3.3 Statistical Analyses 73
4.3.4 Community Involvement 74
4.4 Results 74
4.4.1 Levels of Physical and Emotional Connectedness and Loneliness 74
4.4.2 Relations between Social Connectedness and Loneliness . . . 75
4.5 Discussion 76
4.5.1 Limitations and Future Research 78
4.6 Conclusions 80

5 SiamCircle: Trajectory Representation Learning in Free Settings 83
5.1 Introduction 84
5.2 Related Work 87
5.3 Problem Definition 88
5.4 Methodology 88
5.4.1 Trajectory Augmentation 89
5.4.2 Siamese Network Architecture 90
5.4.3 Loss Function: Circle Loss 90
5.5 Experiments 91
5.6 Results 94
5.7 Conclusion 99

6 Modeling Group Behaviour via Spatio-temporal Data 101
6.1 Introduction 102
6.2 Related Work 104
6.2.1 Dyad-based Modeling 105
6.2.2 Context-based Modeling 105
6.3 Problem Formulation 106
6.4 Methodology 107
6.4.1 WavenetNRI Framework 107
6.4.2 T-DANTE Framework 110
6.5 Experiments 111

6.5.1 Datasets . . . 111
6.5.2 Evaluation Metrics . . . 112
6.5.3 Baselines . . . 113
6.5.4 Implementation Details . . . 114
6.5.5 Results . . . 115
6.5.5.1 Opentraj Datasets . . . 115
6.5.5.2 Spring simulation datasets . . . 115
6.6 Conclusion . . . 116

7 General Discussion and Conclusion **119**
7.1 Summary . . . 120
7.2 General Discussion . . . 121
7.2.1 Multiple Environments . . . 121
7.2.2 Individual Experiences . . . 123
7.2.3 Spatio-temporal Dynamics . . . 124
7.3 Limitations and possible directions for future works . . . 126
7.4 Conclusion . . . 128

8 Appendix **131**

References **137**

Acknowledgments **155**

About the Author **157**

CHAPTER 1

General Introduction

1.1 Background

The rapid growth of technology in wearable devices, such as smartphones and smartwatches, has created huge amounts of data. The availability of wearable data has impacted our daily lives in different ways. We might not have been interested in physiological metrics such as step counts or heart rate a few years ago, but now, many people regularly check health applications on their smartphones to track these measures. We have become our personal data experts, looking at a bar chart and figuring out why we walked more some days than others. Yet, the use of data is not limited only to personal applications. The availability of wearable data creates various possibilities for us and society to address fundamental challenges.

This is particularly evident in micro-communities, i.e., small, localized groups of individuals with common interests, such as athletes in sports clubs, the elderly in nursing homes, and children in playgrounds. In sports clubs, for example, athletes use wearable technology to analyze their sports resilience [10]. In nursing homes, wearable data captures elderly conditions, such as detecting fall behavior [11]. Similarly, children's physical activity levels in playgrounds can be registered through wearable data to assess their physical health condition [12].

Besides individual physical health applications, micro-communities might also be beneficial in other areas. From a psychological point of view, micro-communities provide opportunities for socialization, forming friendships, and feeling a sense of belonging [13–15]. Yet, micro-communities may include barriers that hinder individuals' social and physical development [16–18]. Thus, collecting and processing data from these communities could play a key role in addressing these barriers and creating a positive environment that promotes equity among all children.

From the perspective of data science, the challenge is how to effectively collect, process, and interpret data from individuals in micro-communities as objectively and comprehensively as possible and derive meaningful insights. The obtained knowledge can facilitate various stakeholders, such as schools, psychologists, architects, and policymakers, to address fundamental challenges in micro-communities.

1.1.1 Children in Playgrounds

Children in playgrounds, as an instance of a micro-community, is particularly interesting. In playgrounds, children develop their social skills through play and socializing [19, 20], and enhance their physical skills by engaging in different active games [21–23]. Yet, playgrounds may contain unexpected barriers that hinder children's social and physical development, particularly for those with communication difficulties, such as children with cochlear implants or autistic children. For instance, several studies reported higher sensory sensitivities perceived by autistic children [24–26]. This higher level of sensitivity can increase discomfort in autistic children in noisy and unstructured environments. As a result, crowded areas around popular play structures in playgrounds might overwhelm autistic children and exclude them from interacting with their peer groups. This exclusion can result in loneliness, which ultimately affects children's quality of life [27–29]. Playgrounds should become environments where all children feel welcome, accepted, and included.

Creating such an environment starts with understanding children's capacities and needs, on the one hand, and identifying playgrounds' limitations and possibilities, on the other hand. Hereby, the challenge becomes how to collect, process, and interpret data from children (on an interpersonal level) and their interactions with playground environments (on environmental levels) to obtain this understanding as accurately as possible. Our understanding of playground dynamics, including interactions with the design features, peer groups, and cultural constraints, provides valuable knowledge for various stakeholders, such as schools, designers, and policymakers, to address fundamental challenges and limitations in playgrounds.

1.2 Playground Analysis Framework

To achieve this, there is a need to develop a data analysis framework that can examine individuals' behaviors at interpersonal (i.e., children) and environmental (i.e., playground) levels. This data analysis framework addresses differences in individual needs in playgrounds, which serves as a compass for equity in the playground context. To design such a framework, we first need to understand the characteristics of playgrounds, especially those that might introduce unique challenges to our data analysis framework. For this purpose, the three characteristics of playgrounds are identified as follows:

- *Multiple Environments.* Children are not isolated in playgrounds. Instead, they constantly interact with multiple environments, including the physical design (e.g., play structures, layout, and materials in playgrounds), peer groups (e.g., interactions with peers), and local rules and constraints (e.g., restrictions on the use of certain areas). Our framework must comprehensively and unobtrusively incorporate these environments surrounding children.

- *Individual Experiences.* Capturing children's actual behaviors in playgrounds and identifying their social networks does not tell us how they feel or what they want. For example, as an outsider, we might interpret a child's experience who is sitting on their own differently from what that child is experiencing.

Thus, another crucial element is understanding individual differences in their personal perspectives and experiences. These individual differences must accurately and unbiasedly be included in our framework.

- *Spatio-temporal Dynamics.* Children often participate in various activities across different areas in playgrounds. For example, a child might spend the first few minutes playing with swings, then join a group to play soccer, and then walk on their own in the last minutes of recess time. A record of these dynamic changes, both spatially and temporally, should be consistently assessed in our framework to make an accurate picture of a playground.

To address these unique characteristics, our analysis framework should include a data acquisition tool that captures individuals' interactions within multiple environments in playgrounds over time and space, coupled with their experiences. Additionally, the analysis framework should be able to process and interpret such multidimensional data to create understandings and insights of limitations and possibilities in playgrounds. Providing this knowledge enables us to propose practical interventions to address the limitations and promote equity for all children.

1.3 This Thesis

The main goal of this thesis is to develop a framework for analyzing children's behavior in playgrounds. The proposed framework has addressed the three main characteristics of playgrounds, i.e., multiple environments, individual experiences, and spatio-temporal dynamics, as depicted in Figure 1.1, as follows:

Feature 1. Multiple Environments.

The primary step in designing the analysis framework is identifying environments that may impact children's behavior and experiences. This exploratory phase enables us to define essential variables for capturing the identified environments and choose proper data acquisition tools to measure the defined variables. In fact, the foundation of the analysis framework lies in a data acquisition methodology to comprehensively, unobtrusively, and consistently collect data from children's behavior in multiple environments. Previous studies often relied on qualitative data such as observations and self-reports [30–32]. However, these methods are not capable of effectively capturing all contributing elements over recess time. To address this, some studies have utilized sensing technologies, e.g., wearables, to examine one specific behavior, such as examining physical activity levels [33, 34] or face-to-face contacts [7, 35, 36]. Yet, a gap remains in designing a data-driven approach based on sensors to collectively capture children's behavior in all related environments

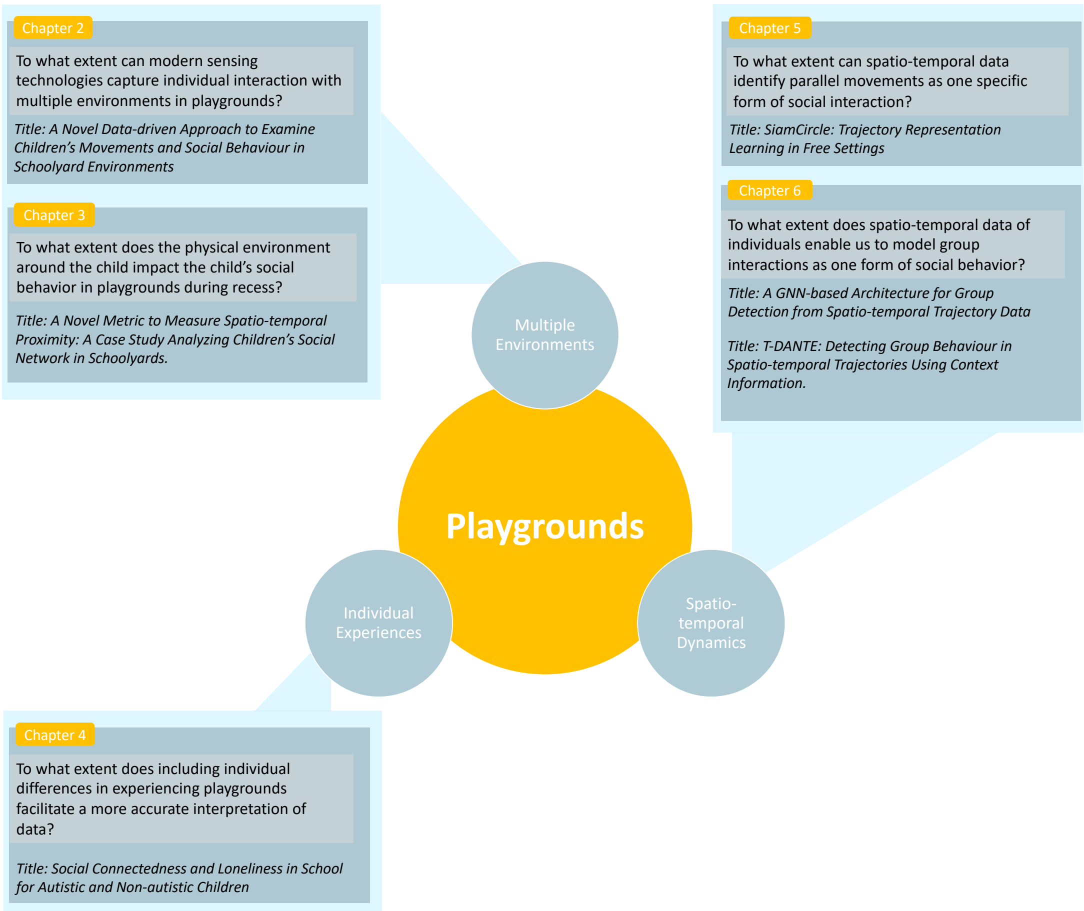


Figure 1.1: An overview of characteristics of playgrounds including (1) Interconnected Environments, (2) Individual Experiences, and (3) Spatio-Temporal Dynamics, that pose challenges to our data analysis framework.

surrounding them over recess time in playgrounds. Therefore, the first research question of this dissertation is:

RQ. 1. (Chapter 2): To what extent can modern sensing technologies capture an individual's behavior in playgrounds?

This chapter introduced a novel sensor data-driven approach for capturing children's behaviors. Furthermore, analyzing the sensor data enabled us to identify the environmental aspects the child interacts with in playgrounds.

The next step is quantifying and measuring the impact of these identified environments on children's behavior. For example, installing various play structures in playgrounds might impact children's social behavior differently. One might stay in one crowded area around a popular play structure and interact with peers, while the other might mingle around different spots and make interactions. How can we measure such micro-behavior in the data and analyze their impact over time and space? This has shaped the second research question as follows:

RQ. 2. (Chapter 3): To what extent does the physical environment around the child impact the child's social behavior in playgrounds during recess?

This study explores the impact of spatial features, such as play structures and physical features in playgrounds, on children's interactions with peers. The goals of this study are (1) to introduce a novel spatio-temporal metric to facilitate this investigation and (2) to apply this metric in a case study to obtain a comprehensive understanding of children's social networks in playgrounds.

Feature 2. Individual Experiences.
Including individual experiences enriches the psychological understanding of children's behavior in playgrounds. While sensor data provides multimodal information about children, examining how these measures are aligned with self-report data is crucial. For instance, if sensor data indicates frequent interactions between a child and their peers, does this also suggest that the child experiences a lowered level of loneliness? Thus, the third research question is formulated as follows:

RQ. 3. (Chapter 4): To what extent does including individual differences in experiencing playgrounds facilitate a more accurate interpretation of data?

This chapter aims to include individual experiences, in the form of a self-report, to the sensor measurements to create a clearer picture of children's behavior in playgrounds. Specifically, we use the loneliness measure via self-reports, peer acceptance via peer nomination report, and sensor data to see how children's connectedness in sensor data relates to their level of loneliness.

Feature 3. Spatio-temporal Dynamics.
Analyzing spatio-temporal dynamics enables us to understand an individual's social behavior in space and over time. Previous chapters focused only on face-to-face contacts in playground datasets to capture social interactions, discarding interactions that might happen in other settings, e.g., parallel play. In the next two chapters, we focus on capturing more complex interaction patterns in group settings using artificial intelligence (AI). Due to the necessity of access to ground truth data (i.e., group membership information of individuals over time) to design, train, and evaluate these models, all studies experimented on the Opentraj dataset [37]. This dataset has been collected from pedestrian movements in constrained environments, e.g., university campuses, which to some extent, is comparable to children's movements in playgrounds. In Chapter 5, we are interested in interactions occurring in parallel play, such as running and walking side by side. We have addressed this gap in the next research question:

RQ. 4. (Chapter 5): To what extent can spatio-temporal data identify parallel movements as one specific form of social interaction?

In order to address this research question, we incorporated an AI model to capture similar movement trajectories. Such an AI model provides valuable insights into children's social behavior during play by identifying parallel movements in the playground.

Yet, human social behavior includes complex spatio-temporal patterns that can not be easily modeled via pre-defined forms of interactions such as face-to-face contacts or parallel movements. For example, children's interactions during hide and seek do not include these forms of interactions. Thus, it is important to include both spatial and temporal elements of human interactions in the design of the AI model to capture children's complex social behavior in playgrounds. We specifically addressed this challenge in our next research question as follows:

RQ. 5. (Chapter 6): To what extent does spatio-temporal data of individuals enable us to model group interactions as one form of social behavior?

Chapter 6 aims to adopt AI models to analyze spatio-temporal features in movement trajectories and identify group interactions. By conducting two parallel studies, we examined the use of spatio-temporal data in the form of dyad interactions as well as the inclusion of context information alongside dyad interactions to model group behavior.

Overall, our analysis framework thoroughly addresses the unique characteristics of playgrounds by combining cutting-edge technological advancements with multidisciplinary collaboration involving psychology, architecture, and computer science experts. Implementing this framework offers a valuable tool for schools, sports clubs, policymakers, and designers, enabling them to identify individual and group-level challenges and barriers in micro-communities.

1.3.1 Organization of Thesis

The organization of the thesis is illustrated in Figure 1.1. In Chapter 2, we introduce our sensor data-driven approach that can be used for collecting spatio-temporal data from children in the playground. Introducing a new metric to measure spatio-temporal proximity is discussed in Chapter 3. Chapter 4 discusses the inclusion of children's experiences in our framework. Chapter 5 presents an AI model to capture similar movements. In Chapter 6, we tackle the problem of identifying group behavior via spatio-temporal data. Finally, several remarks are presented in Chapter 7 to conclude this thesis.

CHAPTER 2

A Novel Data-driven Approach to Examine Children's Movements and Social Behaviour in Schoolyard Environments

The contents of this chapter are based on the following publication:

- M. Nasri, Y.-T. Tsou, A. Koutamanis, M. Baratchi, S. Giest, D. Reidsma, and C. Rieffe, "A novel data-driven approach to examine children's movements and social behaviour in schoolyard environments," *Children*, vol. 9, no. 8, p. 1177, 2022. DOI: 10.3390/children9081177.

Abstract

Social participation in schoolyards is crucial for children's development. Yet, schoolyard environments contain features that can hinder children's social participation. In this paper, we empirically examine schoolyards to identify existing obstacles. Traditionally, this type of study requires huge amounts of detailed information about children in a given environment. Collecting such data is exceedingly difficult and expensive. In this study, we present a novel sensor data-driven approach for gathering this information and examining the effect of schoolyard environments on children's behaviours in light of schoolyard affordances and individual effectivities. Sensor data is collected from 150 children at two primary schools, using location trackers, proximity tags, and Multi-Motion receivers to measure locations, face-to-face contacts, and activities. Results show strong potential for this data-driven approach, as it allows collecting data from individuals and their interactions with schoolyard environments, examining the triad of physical, social, and cultural affordances in schoolyards, and identifying factors that significantly impact children's behaviours. Based on this approach, we further obtain better knowledge on the impact of these factors and identify limitations in schoolyard designs, which can inform schools, designers, and policymakers about current problems and practical solutions.

2.1 Introduction

Children spend a considerable amount of time in schoolyards, engaging in loosely structured activities under relatively mild supervision. The schoolyard environment, therefore, presents unique opportunities for children to play and develop their physical and social skills. Unfortunately, schoolyards may also pose unexpected obstacles that limit social play in various ways. For example, poor acoustics might hinder children who face barriers to communication, including children with hearing loss or autistic children. Examining children's behaviour in the context of the particular environment where it occurs could aid in identifying existing limitations and possibilities for schoolyard design. It could also inform the development of methods for improving schoolyards, and make it possible to tailor designs for sensitivity to differences in children's needs, desires, and capacities. This, in turn, could help maximise opportunities for social learning for all children, including those who belong to vulnerable populations.

However, making this picture clearer requires a huge amount of data: precise information gathered over time, pertaining to not only different aspects of children's behaviour, but also to the specific environment in which the behaviour occurs. This study, therefore, presents a sensor data-driven approach for gathering the information necessary for examining the effects of the schoolyard environment on children's movements and social behaviours.

2.1.1 Affordances and Effectivities in Physical, Social, and Cultural Environments in Schoolyards

When children are in a schoolyard, they are constantly confronted with at least three layers of the schoolyard environment: the physical/built environment (physical layout and features) [38, 39], the social environment (people to interact with) [40, 41], and the cultural environment (rules and constraints set by schools) [42–44]. All three layers present different affordances to the children (see Box 2.1.1). Affordances are the actionable properties an environment presents to a child (e.g., a sand-pit affords building a sandcastle), in relation to the children's individual desires, needs, and capacities [45]. That is, an environment's affordances are relative to specific actions. For example, a sand-pit that is empty (without any sand) makes no difference in affording opportunities to a child who wants to be running around. Yet, it stops those who want to build a sandcastle from doing so. Certain interactions require an appropriate setting: a quiet, secluded corner for confidential talks, or a wide-open area for a large game involving physical activity. Certain settings stimulate certain activities and behaviours, as one can observe around any piece of schoolyard equipment; and certain activities are subject to school rules and conditions, e.g., football or cycling may be permitted only at specific places and in particular times.

Several studies incorporated a perspective on affordances to focus on interactions of children in general with their physical, social, and/or cultural environment. Many affordance studies concerning children, schoolyards, and schools depart from Heft's categorizations [45]. Heft distinguished between ten types of outdoor environments, such as "flat, relatively smooth surfaces" (which may afford walking, running, cycling, skating, skateboarding) or "attached objects" (which may afford sitting-on, jumping-on/over/down-from), and further extends affordances to include social and emotional behaviours [46, 47]. Physical affordances are mostly studied in children-oriented research, including studies on how different environments (e.g., home, school, sport, leisure, neighbourhood, outdoor play) either promote or hinder various motor activities [48–50], and how physical, social, and cultural affordances influence physical activity levels in schoolyards [51, 52].

Box 2.1.1. Affordances in schoolyards

For the purposes of our research, we distinguish between three levels of affordances:

- Physical affordances: what the physical layout and features of the schoolyard afford to children and their activities. These are critical for many vulnerable children, to the extent that they may even exclude themselves from what takes place in the schoolyard. For example, what most humans tolerate as mild background noise can be insufferable to children with cochlear implants, who, consequently, tend to refrain from entering schoolyard areas where exposed to such noise [53, 54].
- Social affordances: These refer to two complementary matters:
  - What features in the schoolyard afford social interaction, i.e., social interactions in our case, should be accommodated and facilitated by the environment. For example, having a chat with a classmate requires some sitting furniture in a quiet part of the schoolyard. This involves not only the need for a suitable environment for social interactions but also the features in the environment that stimulate social interactions (such as the presence of a seesaw, which invites play with another child).
  - How the presence of others adds to or detracts from the affordances of the physical environment. For example, if a swing is already occupied by another person, then the child is unable to sit on it. However, a new affordance becomes available: for example, pushing the person sitting on the swing.
- Cultural affordances: Free play and schoolyard use are normally subject to constraints, where, for example, some intensive or hazardous activities (such as football or cycling) are allowed only in certain parts of the schoolyard, or for a specific period of time.

Yet, when studying how schoolyard environments afford opportunities for children to play, it is important to consider all three layers of the environment: physical, social, and cultural. These are closely intertwined, and ignoring any layer could bias any data analyses and interpretations. For example, if a play area is too small, children who come late to the game may be excluded simply because there is no room for them. Such an outcome due to capacity issues may not necessarily reflect social exclusion. However, this example illustrates restrictions imposed by the design and operation of a schoolyard. Moreover, if the social environment is not taken into account when considering the physical and cultural characteristics of a particular schoolyard, then schools, designers, and policymakers could be kept unaware of the limitations and possibilities of that schoolyard. Consequently, opportunities for improvement could be missed.

Of particular interest are vulnerable children (e.g., children with a clinical diagnosis or disability) who might have different desires, needs, and capacities ("effectivities") in their use of space as compared to other children in the same environment. For example, autistic children may be sensitive to certain ambient triggers (sounds, light, or touch) or avoid being in crowded areas [55–57]. They often prefer repetitive games with predictable results, such as spinning, twirling, and illuminating [58, 59] and fixed routines, with clear instructions and rules to follow [60]. Autistic children can also find initiating or maintaining social contact with other peers quite challenging [61]. Children with attention deficit hyperactivity disorder (ADHD) are observed to often change activities during break time, and many ADHD children have difficulties sustaining interactions with peers [62–64]. Thus, for vulnerable children, it may be especially critical to unravel the relationship between their individual interactions and their environment.

By identifying affordances, we can be explicit and transparent in two critical aspects that are especially relevant to the inclusion of vulnerable children in schoolyards. First, working through a lens of affordances makes an explicit definition about vulnerable children's capacity, so as to know what they expect, want or can do. This, consequently, facilitates awareness of special needs. This is notable because although special needs are usually considered in the teaching activities in the classroom, they may be ignored in the design and use of the schoolyard. Second, taking affordances into consideration clarifies the influences of the physical environment on children's activities. Consequently, making affordances for vulnerable children explicit can help identify existing limitations, and develop methods for the analysis and evaluation of school environments.

Taken together, we see a need to understand individual interactions at the micro-level of peripersonal space, while still taking into account the physical, social and cultural layers of the children's direct environment, and the challenges they may pose for vulnerable children. To the best of our knowledge, there are no prior studies that examine how the environmental triad of physical, social, and cultural affordances interact with children's behaviours and movements in the schoolyard. Moreover, in available affordance studies, vulnerable children have never been considered. While a large body of literature has reported on vulnerable children's physical activity levels, forms of play, and social connectedness in the schoolyard [29,65–67], outcomes that were reported were not linked to any environmental factors.

2.1.2 Present Study

How do we identify a schoolyard's affordances in relation to the effectivities of the children who use it? Most of the previous affordance studies relied on qualitative data, such as observations and self-reports. Although informative, these methods might not examine different interconnected aspects in a cost-effective way, nor give the detailed level of information necessary to draw reliable conclusions. Yet these objectives might be achieved by using the newly available sensor technologies. Sensor data promise comprehensive coverage of what takes place in a schoolyard at a low cost. They make continuous, objective monitoring of activities and interactions feasible. They provide reliable reports on schoolyard performance and enable schools to identify problems as soon as they emerge. Some recent studies applied Global Positioning System (GPS) trackers and accelerometers. This new data-driven approach has been used, for example, to examine physical activity levels [68], and to compare active outdoor play in schoolyards and in natural environments, taking into account personal characteristics (e.g., age) as well as the physical and social environment [69–72].

The main challenge is to collect such precise information from different layers and unravel the complex relationship between environmental affordances in schoolyards and children's effectivities. To deal with this, we designed a data-driven approach for collecting data that would feature enough detail and precision to inform us about relations among three different environmental layers (physical, social, and cultural) and the children's role in these. We, then, examined the extent to which children's movements and social behaviours were affected by the physical, social, and cultural affordances of a schoolyard by identifying three successive aims.

First, we aimed to develop a novel sensor data-driven approach by integrating unobtrusive data collection. This technology included GPS loggers to obtain children's location, their trajectory and speed of movements, Bluetooth-based proximity tags to examine face-to-face contacts of individuals, and multi-motion receivers (MMR) to obtain the physical activity level of children. This approach enabled the monitoring of children's activities in the schoolyard, their contacts with peers, and their movements within the environment during unstructured breaks at school. Multimodal analyses of sensor data yielded a detailed, precise picture of children's interactions with peers and their direct environment. Data and results obtained through these new methods based on sensor data were validated using video observations of these schoolyard events.

Second, we aimed to distinguish between three interconnected types of affordances (physical, social, and cultural) and gave each of these explicit and measurable definitions (see Box 2.1.1). By operationalizing these terms, the data could be interpreted with greater precision and less bias. To illustrate the value that these data can have, we analyzed data collected from two schoolyards. Integrating the sensor data based on the triad of affordances, in addition to providing extensive information on each individual aspect, provided extensive interdisciplinary knowledge on how the physical environmental features affected children's social participation and movements. In addition, it highlighted how the presence of other individuals affected children's behavior and movement in the physical space. It also revealed how the rules set by schools and supervisors affected children's social participation and use of space.

Third, by considering the relevance of these data, we aimed to better understand how the collected data and their analyses could inform schools, designers, and policymakers about the possibilities and limitations a schoolyard presents, and plan for practical solutions and improvements, particularly with respect to the individual differences in effectivities of vulnerable children.

2.2 Methodology

Our methodology addressed two main goals simultaneously. On the one hand, we developed a setup of sensors to gather data in the schoolyards, and an approach to work with these, in the context of two schools. On the other hand, we also carried out a simultaneous study where we gathered data from children in these schoolyards and analysed them to gain insight on children's behaviour, and their relation to the three environmental layers (physical, social and cultural). This section presents this integrated methodology.

2.2.1 Selection

We developed our data-driven approach and applied our approach in the context of two primary special-needs schools that were geographically located in the centre of the Netherlands. The schools and parents were informed about the purpose and planning of the study, and supplied written consent for the children to participate in it. Approval for the study was obtained from the Leiden University Ethical

Committee. The data-management procedures were registered and approved by the Leiden University Research Data Management Plan.

All sensor belts were prepared and handed over to the teachers in a box 15 minutes before the break. Due to the COVID-19 restrictions, the examiners were not allowed to be in the classrooms to help children with their sensors. Instead, prior to data collection, a video presentation was shown to children that explained the research in simple words, to prepare them for participating in this study and instruct them on the use of sensor belts during the break. Teachers further helped children to put on the sensors since they were fully aware of the instructions for the sensors and each child's specific preferences or capacity. During this process or during the break, children could refuse to wear the sensor belt, if they were not comfortable with it, which occurred in 2% of the breaks among 1–2 children.

2.2.2 Reconnaissance

Prior to data collection, the researchers visited each school for a reconnaissance visit (i.e., to explore the situation with an aim to define a strategy) for investigating the physical, social, and cultural environment. This first contact gave them the opportunity to familiarise themselves with the schoolyard, explore its environmental features, and conduct informal interviews with the school director, teachers, and caretaker during a tour of the school building and its surroundings. This go-along approach is common at the exploratory stage of similar research [72, 73].

Interviews included questions about school customs, teacher and pupil preferences, habits, the organisation of breaks, and schoolyard activities. In combination with the visual inspection of the schoolyard, these provided an initial impression of physical, social, and cultural affordances, and led to hypotheses about the social and cultural context of school breaks, as well as to the selection of locations where sensor facilities should be positioned. Ultimately, the reconnaissance visit informed us about: (1) the proper locations for installing the sensor equipment and for video observers, (2) general rules about breaks (e.g., that children were not allowed to stay in classrooms except on special occasions), (3) and general rules on the use of the schoolyard, such as soft boundaries and area allocation during breaks.

2.2.3 Participants

A total of 150 children aged between 5 and 15 years old participated, from 21 different classes from two schools (schools A and B). Data collection took place over a period of two weeks at each school, respectively, split between two recess times on consecutive days for each class. Each measurement lasted between 15-50 minutes, depending on the playgroup (younger children usually have longer breaks than older children).

Many students at school A came from mainstream education, without a specific diagnosis, because they needed extra care and support and their well-being was often under pressure due to learning pace, large-size classes, or overwhelming contact with others. The school, therefore, offers more structure, predictability, personal attention, and specialist support to improve their well-being. The majority of pupils were undiagnosed, or their diagnoses were unknown to us (63%). Of the rest, most had ADHD (20%) and autism (14%) as their primary or secondary diagnosis. In total, fifteen measurements were conducted in seven days in the period of two weeks.

This school is located in an urban residential area where streets abutted the school on three sides and the backyards of single-family homes abutted school property on the fourth side. Hard borders (e.g., fences or walls) separated the school area from its neighbours. As shown in Figure 2.1, schoolyard use is separated into two parts, with junior classes being allocated a different part (sub-areas I, II, and III) from that of the senior classes (sub-areas IV and V).

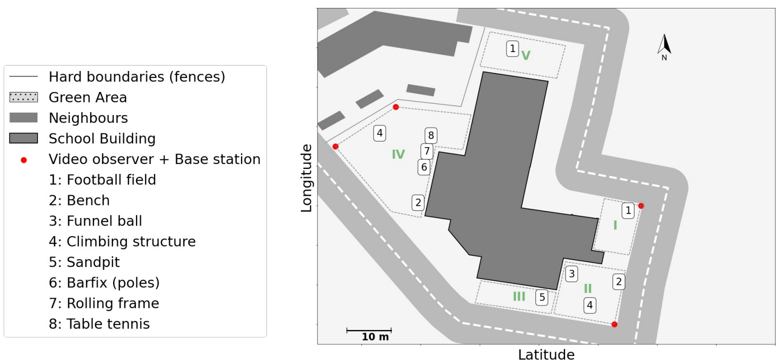


Figure 2.1: Layout of school A.

School B offers education to children whose development is disrupted or at risk of disruption due to reasons such as behavioural problems, emotional problems, or psychiatric issues. Seventy-six percent of students were autistic, and 34% with ADHD as their primary or secondary diagnosis. Therefore, these two conditions accounted for the majority of students in School B.

The school is located in a rural area, in close proximity to green spaces. It is located on the site of a larger complex of special-needs facilities. In the first part of our study (Figure 2.2a), the school shared some outdoor areas of the complex, notably a football field. On the south side, the school bordered residential properties with green areas in between. The schoolyard was therefore demarcated by soft borders on practically all sides. During our study, the schoolyard was renovated. The new layout included a harder yet penetrable separation from the complex, and a higher degree of self-sufficiency, primarily thanks to its own football field (Figure 2.2b). Data collection was conducted in three waves: (1) before the renovation, (2) after renovation, and (3) following minor, local improvements in the renovated schoolyard (6 months after the renovation). In total, eighteen measurements were conducted in six days in the period of two to three weeks per data collection wave.

Figure 2.2: Layout of school B: a) before renovation, b) after renovation.

2.2.4 Validation of Measures Obtained through Sensor Techniques

Two video recorders were present to supply validation for the data collected during the breaks using sensor techniques. Locations for video observers and sensor equipment were determined during reconnaissance visits. In the current study, video recordings were used for visually verifying the sensor data analysis and supporting the observations presented in the result section. Specifically, video recordings were used along with data analysis to ensure that the obtained results and interpretations aligned with what actually happened during the break. For this purpose, all video recordings were stamped with the date and time of the measurements per second. This enabled us to extract a particular period from the sensor data and the corresponding video recordings to verify the results obtained by the data analysis. For example, the presence of children around the ping pong table, the frequent use

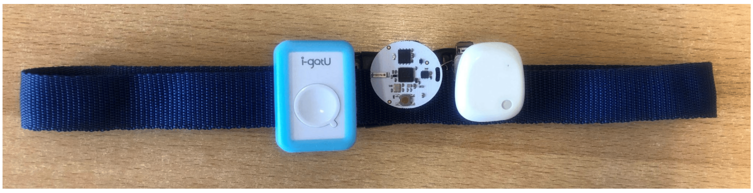


Figure 2.3: The proposed sensing system: GPS logger, proximity tag, MMR sensor (from left to right) mounted on a belt.

of the icy slide, and the popularity of the new multi-functional structure were all verified via the video recordings.

2.2.5 Variables and Measures

As shown in Figure 2.3, the GPS tracker, proximity tags, and MMR sensors were used to measure and analyse children's behaviour in different aspects and their interactions with schoolyard environments. As interactions depend on the context (i.e., the physical and social setting, and user activity in both spatial and temporal dimensions), our study requires multimodal analysis to examine the highly interconnected and sophisticated layers of physical, social and cultural affordances in schoolyards. In this concept, affordances were used as means to interpret and illuminate the result of data analysis. We further identified the main variables per sensor device that were used in the identification of each affordances layer as follows:

2.2.5.1 GPS Loggers

GPS loggers record the location of the wearer, allowing us to track the movement of each child in the schoolyard, i.e., the trajectories they follow and the places they visit. The GPS loggers used were of the i-gotU GT-120 USB type. Noise in the GPS data was removed by keeping only sequences where at least five successive points were situated within a distance of 10 metres of the schoolyard outline, to account for the positional accuracy of GPS loggers. We excluded data points with unrealistic speed (10 m/s cut-off point) [74, 75]. The remaining data points were used for further analysis. In schoolyard affordances, GPS locations were adopted in two directions:

- Trajectories of children contained the longitude and latitude of movements, through which the speed of movements was calculated (speed = displacements over time).

- A kernel density estimate (KDE) estimated the distribution of GPS locations in a playgroup and assessed the most visited areas.

2.2.5.2 Proximity Tags

Proximity tags used in the research were OpenBeacon, with two base stations (Beagle-Bone Black minicomputer augmented with custom OpenBeacon hardware). Proximity tags registered each other via Bluetooth at a distance of up to 1.5 metres. They wirelessly sent data on these sightings to the base stations, which received signals 4 times per second [36, 76] and registered information broadcast by tags up to 25 metres away. The proximity tags were used to detect face-to-face contacts between subjects during recess. Since most children were involved in active play, their body movements, or interfering objects, such as other individuals passing by and toys, may have interrupted the signal. To compensate for this error, the raw proximity data was interpolated by joining two successive contacts between the same peers, if the time gap between the two contacts was less than a certain threshold (35 seconds in our study) [77]. The obtained variable was defined as follows:

- Spatial contacts were calculated by taking the face-to-face contacts from the proximity tag and fusing it with GPS locations. This gave crucial information on where contacts took place in the schoolyard.

2.2.5.3 MMR sensors

The MMR sensor is a wearable device that includes a BMI160 6-axis Accelerometer and Gyroscope, a BMM150 3-axis Magnetometer, and allows continuous monitoring of activities along three axes. In schoolyard affordances, MMR data is used as the following variable:

- Spatial Activity Level is determined based on cut-off points proposed by Puyau et al. [78], who validated accelerometer-based activity against energy expenditure (EE) in children within a 15-second time frame. We were able to adopt Puyau's setpoints because: (1) The average participants' age was similar to our study (6–16 years old), (2) The activities performed in the validation study were the same as children's activities in the schoolyard (walk, run, free-living activities such as computer games, playing with toys, aerobics, skipping, jump rope, soccer). This variable was date- and time-matched to each 1-second GPS data point to obtain how the activity level was related to environmental features and physical affordances.

Participants wore the three devices mounted on a belt of adjustable length, to be worn around the waist. Subjects were asked to wear the belt only during the break and to take it off when the break ended. Teachers supervising the children during the break were also issued with sensor belts to capture their contacts with pupils. As established at the reconnaissance visit, all children were required to leave the classroom during breaks, except when the weather, sickness or other major problem made it unwise.

2.3 Results

By adopting the above data-driven approach, we effectively carried out extensive data collection in two schoolyards. In this section, we present the results of our data analysis, which examined relations between children's behaviour and environmental characteristics. All analyses were performed in Python 3.6.1, within the Anaconda environment. Geographical data for location identification was extracted from the OpenStreetMap. For all three sensors, time was used as a unique identifier (uid), and the merging of datasets was based on the recorded timestamps.

2.3.1 Physical Affordances

With respect to physical affordances, the data revealed that the availability of equipment and furniture, as well as their condition, could be critical to attracting attention and activity. Figure 2.4 shows the KDE plot of all groups in the (a) morning break and (b) lunch break at school B, around the multi-functional structure that was one of the key new features in the renovated schoolyard. This structure includes one plastic slide (solid white colour in Figure 2.4) and a metal slide (white with dot hatch), as well as stairs, rope and rock climbing, a catwalk area, and a spinner structure. On that particular day, due to the cold weather and low ambient temperature, the metal slide had frozen during the morning break. Since an icy surface has lower friction, the metal slide afforded higher sliding speeds and was, therefore, a popular spot, with a higher traffic density in comparison to the plastic slide in Figure 2.4a. Figure 2.4b shows the situation at the same structure during the lunch break: by that time, the temperature had risen, rendering the metal slide less speedy and therefore less popular. In fact, at that time spatial density around the plastic slide was higher.

This sophisticated, unusual play structure did spark children's curiosity and became a popular spot after it was added to the schoolyard renovation. As the heatmap of young children in Figure 2.5a shows, before the schoolyard was ren-

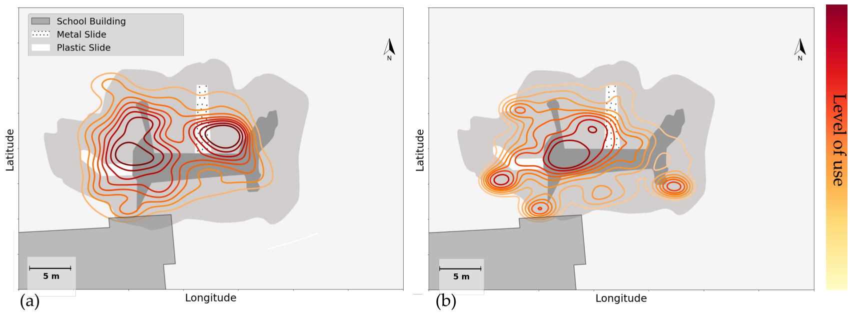


Figure 2.4: Use of space, around the slides on multi-functional structure, analysed by GPS data, during (a) the morning break, and (b) the lunch break in school B. The level of use is analysed via GPS data, and obtained contours are colour coded from the lowest level of use to the highest following the colour bar (from yellow to red).

ovated, the most heated spots were the sandpit and swings (Spots 2 and 3) and the areas around them where children could cycle around. Adding the new multi-structure on Spot 10 attracted children towards this new structure (Figure 2.5b). However, during the follow-up, many children lost interest and reverted to their old preferences (Figure 2.5c; on Spots 2 and 3).

The layout of the schoolyards and the proximity of play structures relative to each other could also affect children's movements and activities. Figure 2.6 shows the physical activity level of children, fused with GPS locations and then mapped to the floorplan. Having a low proportion of vigorous activities in this playgroup suggests that the schoolyard does not offer enough space for high physical activity level activities and games.

2.3.2 Social Affordances

With respect to social affordances, it is equally clear that the density of users affects some activities such as cycling, pushing them away from crowded areas and causing intermittent trajectories. Figure 2.7 depicts the trajectory of a subject in school A, where the child is cycling in the schoolyard between the climbing frame and a yard fence. His speed reaches its highest value in the midway and is reduced near route endpoints and physical structures that provide opportunities for social interactions: the bench where supervising teachers are seated, and the climbing frame where peers are playing. As illustrated in Figure 2.7, near the climbing frame

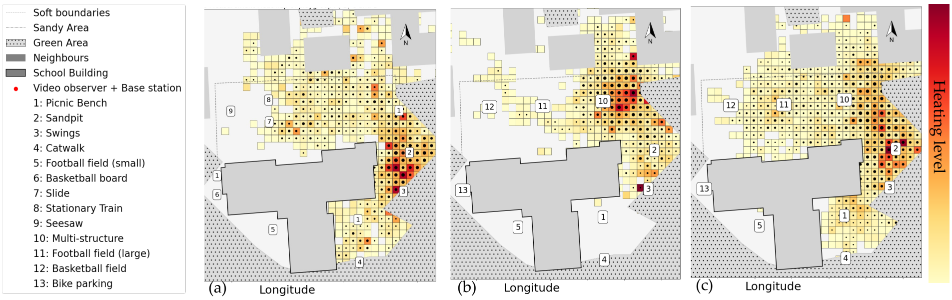


Figure 2.5: Young children's heatmap, analysed via GPS data in school B (a) before the renovation, they were attracted to the newly added structure, (b) after the renovation, and then back to their old habitat during (c) the follow-up. The colour bar shows the duration of the visit. The warmer the colour (following the colour bar from yellow to red), the higher the duration of visits by children. The black dots indicate the number of children in each segment. The larger dots show a higher number of children in that spatial bin.

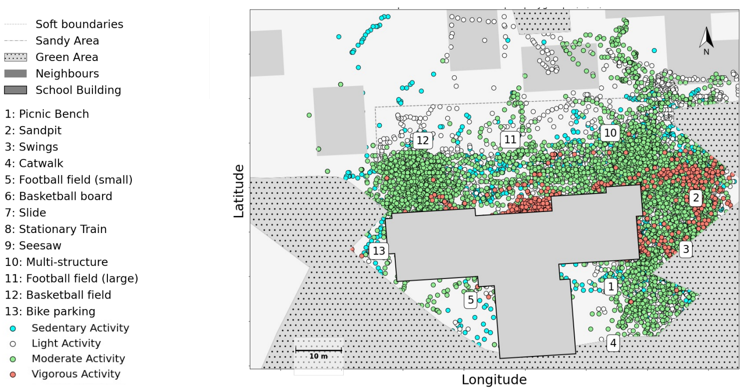


Figure 2.6: The location of physical activity levels in school B was analysed via Accelerometer data and then fused with GPS locations.

the speed not only drops to a minimum, but stationary time also increases (as shown in the red circle: data points with low-speed levels and little displacement along the trajectories around the climbing frame), suggesting the possibility of quick chats with peers. Face-to-face contacts captured by the proximity tags confirm that such social interactions occurred near the climbing frame (green stars) and at the bench (green triangle).

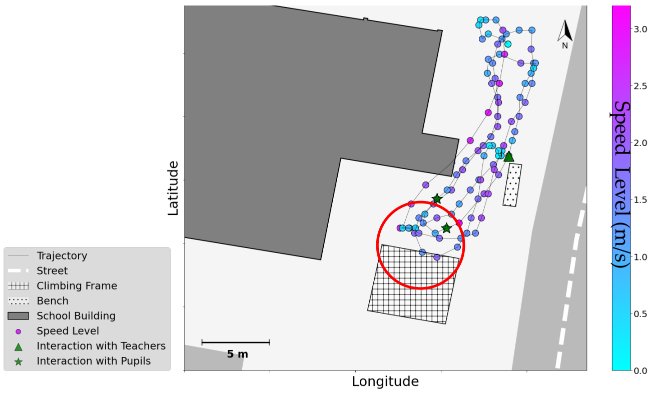


Figure 2.7: Cyclist in school A: the trajectory of movement is extracted from GPS data, and colour coded based on speed level (from light blue to light pink, following the colour bar). Face-to-face contacts are analysed from proximity tags, fused with GPS data, and mapped to the school floor plan (triangle: contacts with the teacher, and star: contacts with peers).

Social affordances could also determine which physical affordances were established. Especially in senior classes, the use of space could be influenced more by who than by what. Children at this age often stay with a specific sub-group throughout the break in a certain schoolyard area that they "own". As Figure 2.8 shows, in school A, the areas around the tennis table and the bench were where most social contacts happened. GPS data confirm the presence of groups from senior classes in these areas. The table was used for sitting and mingling rather than for playing table tennis, which suggests that social affordances are more important than the expected use. Use of the bench by a specific group from senior classes was consistently observed across several break sessions and confirmed by the sensor data (see Figure 2.8).

Data also reveal movement patterns at the individual level, such as the behaviour of an autistic subject in comparison to the rest of the playgroup. Figure 2.9 shows the GPS data of one of the playgroups in School A, where the spatial data of an autistic subject indicates that the child remained close to the school building, avoiding dense areas.

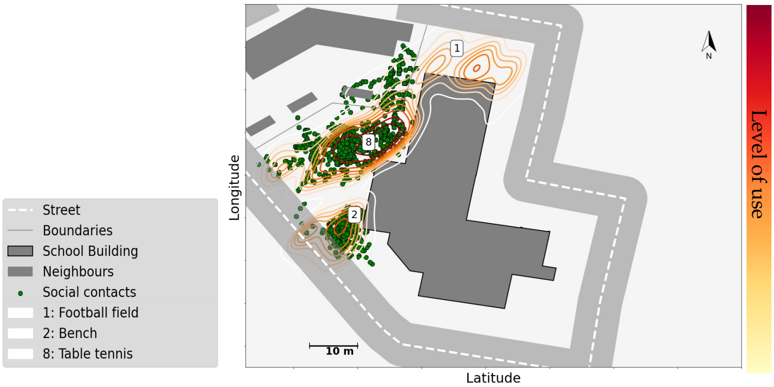


Figure 2.8: Social use of space by senior classes in school A. The level of use is analysed via GPS data, and obtained contours are colour coded from the lowest level of use to the highest following the colour bar (from yellow to red), face-to-face contacts are extracted from proximity data, fused with GPS locations and mapped to the school floor plan (note that the base station could not cover the contacts that occurred in Area 1, therefore no contacts from that region was registered).

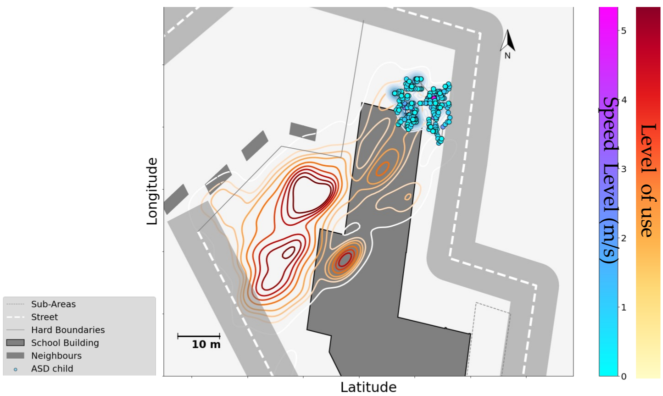


Figure 2.9: The use of space by an autistic child in comparison with its playgroup in school A. The level of use is analysed via GPS data, and obtained contours are colour coded from the lowest level of use to the highest following the colour bar (from yellow to red). The GPS location of the autistic child is plotted on the floor plan, colour coded by the speed level from the lowest speed to the highest, following the colour bar (from blue to pink).

2.3.3 Cultural Affordances

Cultural affordances were quite strong, as expected for this age group and for the capacities of autistic children. For instance, in School A, younger groups were free to use all three sub-areas, I, II, and III. However, wheeled toys (e.g., bikes, steppers, scooters, etc.) were restricted to sub-area III when a specific supervisor was in charge (Figure 2.10a). Since biking was one of the most popular activities among young children, this restriction has a significant impact on the use of space in the schoolyard. Figure 2.10a shows the KDE plot when wheeled play was restricted to III: a high-density level around area III, with a peak at the end of III where the turning point was. Figure 2.10b shows a different day when children were not restricted to III for biking. On that day, the spatial density was widely distributed over all three sub-areas.

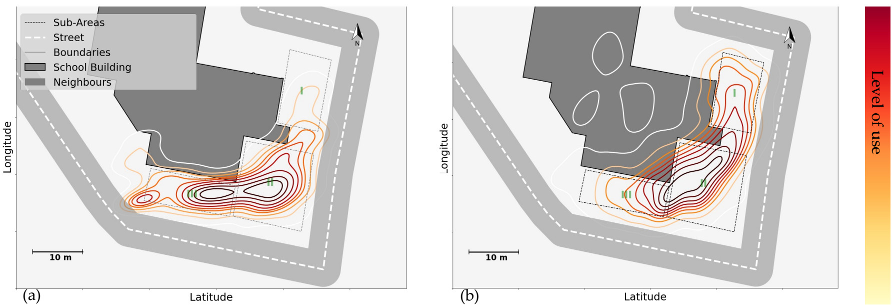


Figure 2.10: The use of space by cyclist children in school A (a) with restriction (b) without restriction from the break supervisor. The level of use is analysed via GPS data, and obtained contours are colour coded from the lowest level of use to the highest following the colour bar (from yellow to red).

Violation of school rules was only occasionally observed, for example, in the trajectories of a few subjects who wandered around in School B before renovation (green dots in Figure 2.11). This contrasted with the trajectories of subjects who went to play on the football field: these followed the shortest route to the remote football field, with no subjects wandering off (purple dots), according to the school rules: children were not allowed to cross the soft boundaries and move around in the residential area.

Supervision naturally reinforced cultural constraints. In fact, in both schools, most constraints were designed with supervision in mind: how to make it more economical and more effective. This, however, also created an illusion of full control

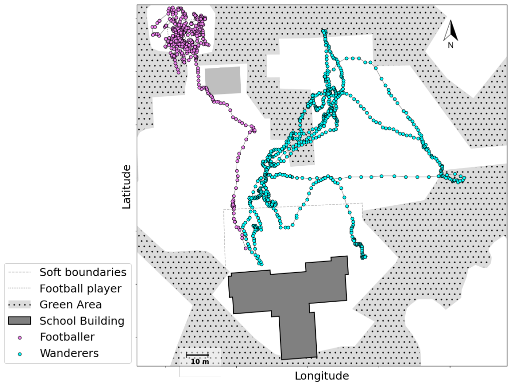


Figure 2.11: The trajectory of a group wandering around against school rules, and football players in school B, was extracted from GPS data.

among the teachers, who were surprised when the researchers reported the above example: they had never observed or even suspected something like that. Otherwise, they would have taken measures to prevent it.

2.4 Discussion & Conclusions

The purpose of this study was to show a data-driven approach that examined how three environmental layers (physical, social, and cultural) interact with children's movements and behaviour during unstructured play at recess time. Through modern sensor technologies, sensor data was collected on children's activities in two primary special education schools in the Netherlands. The obtained data was further analysed in light of schoolyard affordances.

Our first aim was to adopt a novel sensor data-driven approach to examine affordances in schoolyards. Observations by our researchers and feedback from the school demonstrated that the belt we had designed with sensors was not distracting for children. Instead, it was exciting to primary school children, who showed eagerness to participate in the research. Moreover, our data-driven approach allowed us to register more subjects over a longer period of time cost-effectively. The sensing system in our approach includes GPS loggers, proximity tags, and MMR sensors to capture different aspects of children's behaviour (i.e., locations, face-to-face contacts, physical activity level). Integration of this information was crucial in our study. The spatial dimensions of the face-to-face contacts and physical activities were obtained by fusing the registered data with GPS locations. This enabled us to understand how the physical characteristics of the schoolyard impact children's contacts, physical activity levels and their use of space.

Regarding our second aim, our approach allowed us to identify three main environmental factors that influence children's behaviours:

First, the physical capacity of the schoolyard, such as its size, shape, equipment (e.g., availability and arrangement) and relevant rules and constraints, serve as preliminary triggers that affect children's behaviours (e.g., the Sliding example in Figure 2.4; physical activity level in Figure 2.6). The schoolyard should have adequate capacity and offer a variety of options for children to play and engage in different activities. In addition, schoolyard equipment, depending on its design (e.g., climbing frames, swings, seesaw, etc.) and arrangement (e.g., materials, height, size, and proximity to other equipment), could either hinder or attract children to the equipment and discourage or encourage play. Earlier research also confirmed that schoolyard size and availability of play equipment, such as sports facilities, recreation areas, surface materials, and greenery elements, could promote children's physical activity level [79, 80]. Green areas are also found to be a contributing factor in promoting children's resilience and reducing their stress levels [81]. Similarly, close proximity between play structures generates more spots for physical activity [72,82]. Yet, our data showed that close proximity between different play equipment in a small space could also result in lower physical activity levels. The overall shape of the schoolyard influences the supervision method and could result in demarcations that reduce the space available to children (e.g., restricted cyclists in Figure 2.10, versus wandering cyclists in Figure 2.11). Importantly, our three-wave data collection in school B shows that new, fancy equipment may not always remain attractive after the novelty wears off. With time, children may still return to equipment that affords a wider variety of creative activities. This again emphasizes the importance of examining the capacity of the schoolyard according to its affordances.

Second, opportunities for social interaction often attract children towards certain spots and motivate children to use the space for social purposes. Our data confirmed that the spots where such opportunities were offered could indeed stimulate social interaction, even when they were not originally designed for that purpose (e.g., the bench and climbing structure in Figure 2.7; the ping pong table in Figure 2.8). This outcome echoes previous findings about the impact of environmental aspects, such that green and natural elements, multi-functional equipment with diverse structures like sand-pits and stairs, (semi-)secluded places, etc., encourage children's social interactions [34, 83–86]. These features usually afford more variety and flexibility in children's play, and may be helpful for the initiation of social activities, or offer a space to play and hang out without interruption or to recover from active play [34,83]. Conversely, popular areas and equipment can also cause more conflicts over available resources. For example, play structures such as swings, slides, seesaws, etc., can be locations that support constructive interactions, as well as foci of competition, irritation, or even bullying.

Third, individual needs could lead to quite different patterns in the use of space. Despite the patterns shown in the above two points, our data showed that vulnerable children (e.g., autistic and ADHD children who have different effectivities), who have different capacities and needs in their interactions with the environment, may use schoolyard affordances of any kind in a unique way. This is observed via sensor data in their trajectory of movements, use of space, and activities during the break, as with the autistic child who remained next to the school building in Figure 2.9, away from the area where most of his/her classmates were playing. With this data-driven approach, we show that this autistic child was not just alone, but alone in a context: in the corner, with limited movements, throughout the break. Such information that links individual patterns to environmental factors, in children's natural setting, is crucial for studying children's behaviour, especially for understanding the needs of vulnerable children. Yet, while our sensing system enabled us to detect these differences, it is also important to identify whether these differences reflect the preferences of the child, difficulties in joining others, or social exclusion.

Our third aim was to inform schools, designers, and policymakers about these identified factors. School organisations and supervisors in schoolyards play a key role in identifying situations where children are overwhelmed or triggered by peers or certain equipment and avert these scenarios by taking the required precautions and overruling the existing climate [87], for example, by planning a timetable for using a popular play structure by different sub-groups of children. School organisations should also deliver extra support, customised rules, and structures suited to vulnerable children, and an inclusive school climate that values diversity and individual differences. For example, the schoolyard could have different sub-areas in different colours and play structures, and children could choose in which colour (or sub-area) they would like to play before starting the break. In this way, supervisors could estimate high-demand areas and, by re-organizing the available resources, try to strike a balance in the use of space. This could deliver substantial benefits, for example, to children overwhelmed by crowded or noisy situations, such as children with hearing aids or autistic children. These findings have important implications for developing interventions that adapt to the environment to promote social participation. They also show the strengths of our proposed sensing system in evaluating the intervention and providing relevant insights to schools, designers, and policymakers.

Overall, the obtained results showed that making affordances explicit and measurable, through the use of approaches such as those presented in this paper, can help to better understand children's behaviours and moves, translate these insights into actionable advice for schools, designers and policymakers and further improve current layouts and forms of organisation, especially with respect to the capacity of vulnerable children.

2.4.1 Limitations and Future Directions

Our proposed sensing system allowed us to register more subjects and activities with higher precision than we could have with observation methods. Yet, while a belt with sensors may be exciting to primary school children, it may not be suitable for adolescents in high school. Besides, regarding the performance of individual sensors, GPS technology showed great promise in obtaining individual positions. Yet, given GPS accuracy (1–10 m), ultra-wideband (UWB) (accuracy of 10 cm) are better suited to the scale and context of schoolyards [88–90]. Such accurate positioning systems could identify contacts between subjects in a more comprehensive range of actions and interactions (e.g., parallel play) than proximity tags, which record only face-to-face contacts [91]. This enhancement could withdraw the proximity tags from the sensing system. In addition, in our current system, the value of contacts (e.g., does a child identified alone on the playground feel lonely or happy to be left alone) and moves remain unknown. For example, while we could truly obtain from sensor data that a child played alone in the sand-pit, we did not know the reason, i.e., their emotions or preferences. Therefore, the sensing system could be improved in future in three ways: first, it should be made suitable for a wider age range; second, it should include a more accurate positioning system, e.g., UWB; and third, it should incorporate a strategy to let participants actively express their emotions and preferences through focus group interviews and/or by providing real-time responses for example via smartwatches.

Another limitation of this study is that the obtained results were based on a small number of measurements at two special education schools. More data collection would be needed to design an automated monitoring system whose conclusions could be generalised to different environments and scenarios. This was not possible due to the COVID-19 crisis, but it is currently one of the major foci of this research team. Finally, our proposed data-driven approach makes it possible to analyse movement, social behaviour and environmental interactions among children at a specific school. While it makes it possible to gain a better understanding of the current challenges children face, it also holds the potential to reach beyond this understanding alone and empower schoolyard designers to define and monitor the effect of incremental improvements to schoolyards, for example, in the form of new equipment or changes in the physical organisation of the schoolyard. This approach could even be used to examine real-time adaptations to the rules and parameters of digital-physical interactive schoolyards. As such, the work presented here is the first step in what could become a long-term research program of a much broader scope.

For future studies that aim to adopt a similar approach, one of the most important lessons is that each case should be studied before collecting the data: all environmental layers (physical, social, cultural and their interrelations) must be considered to form expectations for the data and their analyses. As described in Section 2.2.2, a reconnaissance visit that featured informal interviews and inspections helped us understand what took place in each schoolyard, its specific circumstances, and the intentions of the school (which should be respected but also critically analysed). As the data analyses demonstrated, there were many factors behind observed behaviours and patterns that may go unnoticed, if researchers focus on a single aspect or goal. In other words, context matters. Depending on the research questions at stake, variables for capturing the environmental layers require clear definitions and particular sensing technologies should be chosen accordingly. For example, the positioning technology should be chosen depending on environmental conditions to record the users' location appropriately and thus reliably analyse the child's interaction with the environment. For example, GPS technology is more suitable for large outdoor areas, while UWB is a better option for relatively small areas. Integrating background knowledge obtained during reconnaissance visits and the collected data from sensors allows for a better interpretation of environmental factors that affect children's movements and behaviour.

CHAPTER 3

A Novel Metric to Measure Spatio-temporal Proximity: A Case Study Analyzing Children's Social Network in Schoolyards

The contents of this chapter are based on the following publication:

- M. Nasri, M. Baratchi, Y.-T. Tsou, S. Giest, A. Koutamanis, and C. Rieffe, "A novel metric to measure spatio-temporal proximity: a case study analyzing children's social network in schoolyards," *Applied Network Science*, vol. 8, no. 1, p. 50, 2023 DOI: 10.1007/s41109-023-00571-6.

3

Abstract

The present study aims to infer individuals' social networks from their spatio-temporal behavior acquired via wearable sensors. Previously proposed static network metrics (e.g., centrality measures) cannot capture the complex temporal patterns in dynamic settings (e.g., children's play in a schoolyard). Moreover, existing temporal metrics overlook the spatial context of interactions. This study aims first to introduce a novel metric on social networks in which both temporal and spatial aspects of the network are considered to unravel the spatio-temporal dynamics of human behavior. This metric can be used to understand how individuals utilize space to access their network, and how individuals are accessible by their network. We evaluate the proposed method on real data to show how the proposed metric impacts the performance of a clustering task. Second, this metric is used to interpret interactions in a real-world dataset collected from children playing in a playground. Moreover, by considering spatial features, this metric provides unique knowledge of the spatio-temporal accessibility of individuals in a community, and more clearly captures pairwise accessibility compared with existing temporal metrics. Thus, it can facilitate domain scientists interested in understanding social behavior in the spatio-temporal context. Furthermore, We make our collected dataset publicly available for further research.

3.1 Introduction

Increasing public awareness about the potential impact of social networks on individuals and society has led to an interest in understanding how these networks function, and how individuals are positioned in their social contexts. In an effort to investigate the structure of these networks, a considerable body of research literature has developed around the theme of social network analysis (SNA) [92–95]. The availability of wearable technologies that allow for the collection of fine-grained data from daily activities, advances in data analysis techniques and computational tools, and the need to understand the dynamics of social networks in various fields such as sociology, sports science, and healthcare, have all contributed to increased interest in SNA. SNA involves a collection of tools that adopt graph theory to perform extensive computational techniques that ultimately identify connection patterns among members of a community [96]. For instance, a static representation graph accumulates relationships between individuals in an episode, calculates static metrics such as centrality measures (e.g., closeness centrality [97]), and estimates the role of any given node (or individual) in the static representation of that network.

However, a static representation graph eliminates all temporal details, such as the frequency of events and the time difference between subsequent events, when this temporal information might include important information about the nature of the interactions at stake [98]. It also does not take into account any spatial relationships featured within those interactions.

Regarding omitted temporal detail, Figure 3.1 demonstrates the limitations of existing approaches. Throughout this paper, we will refer back to this running example - inspired by an in vivo social scenario - to illustrate our proposed method. The example data here reflect the play of four children (A, B, C, and D) during recess in a playground that features four distinct play areas: a sandpit, a bench, a combination of green areas, and a walking path, plus the entrance to the playground. In this scenario, Child-A played in the sandpit during the whole break, whilst Child-B first played in the sandpit, then engaged in a conversation with a peer on a bench, and later walked to the entrance of the playground. According to the data source in Figure 3.1-a, Child-A interacted with peers only at the beginning of recess time (i.e., T1, T2, T3) in the sandpit, where most of the peers appear to have been located during the same period, whereas Child-B interacted with peers across the entire recess time (i.e., T1, T4, T6) in different locations (i.e., sandpit, bench, and the entrance). These facts reflect distinctions that are relevant to social network structure. Yet although Child-A and Child-B had different behavioral patterns during recess, their positions in the static graph representation, as shown in Figure 3.1-b, appear the same: the graph shows an equal number of edges (partners) with equal weights (number of interactions), which yields the same representation as in their static metrics, such as centrality measures. This example illustrates how these static metrics are not sensitive to temporal changes, and fail to capture complex dynamics that could be contained within such data.

To tackle this problem, Kostakos [98] proposed a "temporal" graph representation, in which the relationship between nodes and the position of each node was investigated in the temporal context of an entire network, rather than in an aggregated format. Based on this representation, Kostakos [98] introduced new metrics, including average geodesic proximity and average temporal proximity, which consider edge availability over time and take into account possible wait times for each node before meeting the next node. Specifically, these are measures of how quickly an individual is accessible by their network, and how quickly the network is accessible by the individual. Figure 3.1-c does show this temporal representation: in contrast to a static analysis, Child-A and Child-B are positioned differently in this representation. Here, the temporal metrics preserve information about the order and the frequency of events over time.

Yet, even in this temporal representation of a network, one aspect that is still

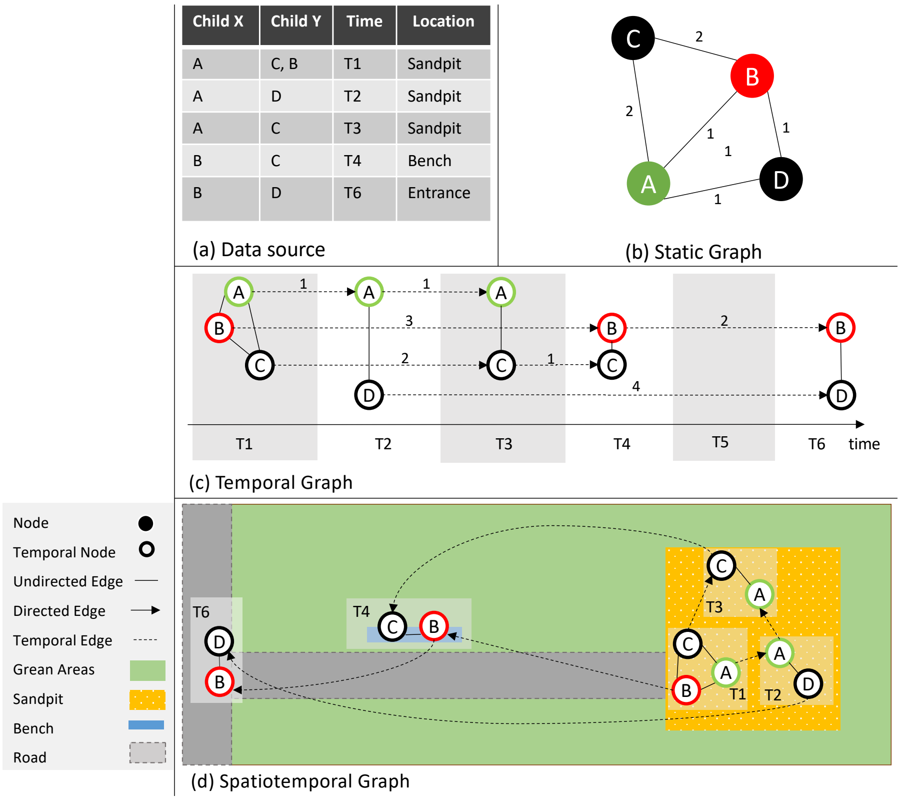


Figure 3.1: Visual comparison on how a data source of interactions between four children in a playground is represented via b static graph, c temporal graph, and d Spatio-temporal graph. While Child-A and Child-B have different temporal patterns (e.g., frequency of interactions and difference in time between interactions) and spatial features (e.g., location of interactions) in the (a) data source, their position in the (b) static graph is the same. The (c) temporal graph could better preserve temporal details, but the spatial information is ignored. And finally, the (d) spatio-temporal representation could preserve the temporal and spatial information of interactions.

entirely overlooked is that of the physical environment. In the context of the example illustrated in Figure 3.1, investigating the temporal accessibility of a child in the playground does not inform us about how the child utilized the environment: did the child walk around and interact with peers (e.g., Child-B in Figure 3.1-d), or simply remain in a crowded spot while making contact with peers (e.g., Child-A playing in the sandpit in Figure 3.1-d)? To understand the impact of the physical environment on an individual's interactions with their network, it is crucial to consider spatial features in analyzing the social network. The temporal metric estimates availability only by considering temporal differences between nodes. The spatial aspect of this availability, i.e., the location of contacts in the physical environment, is still ignored in a temporal graph representation.

The literature has already documented the importance of considering spatial context when examining social networks. Several studies have examined the effect of environments on human behavior, for example, the impact of physical and social environments of neighborhoods on child and family well-being [99], the impact of office design on employee's office usage and social interactions [100], and the impact of schoolyard design on children's movement and their social behavior [1]. In fact, including the spatial element of social networks enables us to understand better how the physical environment impacts individuals' behavior [101] and how users utilize the environment to interact with their network.

While the impact of the physical environment on social networks has been confirmed in various studies, limited attention has been given to the integration of these dimensions into a unified framework. Consequently, there is a critical need for research that addresses this gap by introducing spatio-temporal metrics capable of quantifying the complex interplay between social and physical environments.

In order to examine the interdependency between spatial and social contexts of interactions and continue to build a better understanding of the impact of the physical environment on social networks, the present study pursues two main aims. First, we introduce a new metric for measuring the spatial and temporal dynamics of social networks, known as average spatio-temporal proximity. By considering the spatial differences between temporal networks, our proposed metric examines how spatially sparse these networks are and how spatio-temporally a user is connected to its network. Specifically, this metric calculates the shortest path based on the spatial distance of temporal events and estimates how far, on average, nodes (or individuals) in a network need to move to reach a specific node, and how far, on average, a node needs to move to reach the rest of the network. Such a metric would enable a more comprehensive understanding of the factors influencing social accessibility (i.e., examining how the physical and social context reciprocates around individuals).

Second, we evaluate the proposed metric in the context of a case study whose goal was to investigate social interactions among a group of 32 children playing in a schoolyard during recess. By measuring the spatio-temporal proximity of individuals over time, we aim to better understand the social behavior of these children and their interactions within the context of their spatial environment.

Overall, this paper makes the following contributions:

- We introduce a new metric to quantify the proximity of nodes in a graph that considers both temporal and spatial characteristics of contacts in order to measure how individuals in a network utilize space to make contacts.
- We present and make publicly available a new dataset collected from children's interactions and movements in a schoolyard during recess. This dataset uniquely registered face-to-face contacts and children's locations over time, measured via wearable proximity sensors and GPS loggers, respectively. This dataset is available in a public repository for future use in spatio-temporal research.
- We evaluate the performance of our proposed metric to show how it impacts the performance of a clustering task in identifying clusters in a network, compared to temporal measures alone.
- We further demonstrate how this metric can be used to illuminate the underlying interactions in our Schoolyard Dataset.

The rest of this paper is organized as follows: In Related work, we discuss the literature relevant to our study. The Problem Definition introduces the computational problem at stake in our proposed methodology. Methodology presents the details of our method. The experimental settings and results are presented in Experiments. Finally, Conclusions summarizes the study and discusses future research directions.

3.2 Related work

To provide relevant context, here we reference further literature on two main approaches to the fusion of spatial information with time-series data: location-based social network (LBSN) analysis, where spatial information is intertwined with social networks, and spatio-temporal pattern mining, where relevant research is scanned to identify spatio-temporal features, as a broad concept.

3.2.1 Location-Based Social Network (LBSN) Analysis

Most studies in LBSN analysis arose from the urban computing field and social media networks analysis mainly due to the availability of large-scale databases such as Foursquare, Twitter, Instagram, and Google Places [102–105]. In this body of research, the interdependency between spatial and social contacts lies in the co-presence of two persons in the same physical locations, or the sharing of a similar location history, common behavior, or activities. In these cases, interdependency is implied from users' location data without assessing face-to-face interactions from individuals' perspectives [106]. The aim of this line of research is often to design physical location (or activity) recommendation systems [107–109], travel planning [110, 111], and marketing [112] from accumulated spatial behavior, mainly by considering similarity among users. In contrast, spatial information in the present study is used as a descriptive machine learning model for the purpose of improving the performance of a downstream task and for discovering more about the nature of interactions rather than for predicting social contacts or common interests.

3.2.2 Spatio-temporal Pattern Mining

Recently, interest has been growing in areas related to the adoption of deep neural network models to generate spatio-temporal features in varied applications. These include crowd flow prediction [113], social event detection [3, 114], and financially aware social network analysis [115]. Although the performance of deep models for extracting spatio-temporal features has appeared promising, these models require a huge amount of representative data to learn about generalizable spatio-temporal embeddings. Therefore, these methods are hardly applicable for small datasets in which the duration of data assessments and the number of users is limited. In addition, collecting large-scale data from certain populations (e.g., patients in a hospital or children in a playground) is often impossible because of privacy protection concerns. Therefore, analyzing spatial patterns over time via classical pattern mining techniques has been the focus of many studies in domains such as animal revisitation analysis [116], disease spread patterns [117], and the tourism industry [118, 119]. This line of research often involves extracting time domain features (e.g., residence time, the time between visits) and applying a downstream task to the extracted features, such as clustering and prediction.

Analyzing temporal events via graphs, as in a "temporal graph", was first introduced by Kostakos [98]. Using this representation, temporal metrics were defined to assess geodesic and temporal proximity. These metrics estimated how quickly a node was accessible by the network, and how quickly a network was accessible by a node. Accessibility in temporal proximity was defined as the time difference between temporal events in a weighted shortest-path algorithm. Geodesic proximity was calculated using the number of hops between temporal events in an unweighted shortest-path algorithm. Neither metric took into account the spatial context of temporal events, e.g., the geo-distance between nodes.

The present study is intended to further build upon the work of Kostakos [98]. Specifically, temporal metrics proposed by Kostakos [98] are further developed to incorporate the spatial aspect of social contacts and to examine the spatio-temporal accessibility of individuals within a network.

3.3 Problem Definition

Assume a spatio-temporal proximity dataset as a set of tuples in the form of $\langle v_i, v_j, s_t \rangle$ where $\{v_i, v_j \in \mathbf{V}\}$ are two entities in the set of entities $\mathbf{V}$ that are detected in proximity to each other, where $s_t = (x, y, t)$ represents the coordinate where the contact between these v_i and v_j is registered at time t, and $t \in \{1, \cdots, T\}$ represents the timestamp of the contact. Such a dataset can be collected, for example, when the two entities are equipped with a proximity sensor (to detect contacts) and a GPS sensor, to acquire the average coordinate representing the point of contact between the entities.

First, we are interested in creating a spatio-temporal graph representation $g = \langle \mathbf{V}_{s_t}, \mathbf{E} \rangle$, where $\{v_{i,s_t} \in \mathbf{V}_{s_t}\}$ are sets of nodes each represented over timestamps t, at a location denoted by s, and $\mathbf{E}$ is a set of instant and spatio-temporal edges. Instant edges are unweighted and undirected links between spatio-temporal nodes v_{i,s_t} and v_{j,s_t} indicating contact between two distinct nodes v_i and v_j, $i \neq j$ at location s in time t. Spatio-temporal edges are weighted and directed links between consecutive pairs v_{i,s_t} and $v_{i,s_{t+1}}$ of node v_i (which is weighted based on the spatio-temporal proximity assuming that the direction of edges) represents the temporal order.

Second, we are interested in finding clusters $C_i = \mathbf{V}_{C_i}$ where $\mathbf{V}_{C_i} \subseteq \mathbf{V}$, such that all nodes $\mathbf{V}_{C_i}$ within each cluster C_i are spatio-temporally similar to each other and dissimilar to the nodes in other clusters.

3.4 Methodology

In this section, we present essential mathematical preliminaries as the foundation for understanding our proposed methodology, and the problem at stake. Then, we explain our proposed spatio-temporal graph and its proximity metrics.

3.4.1 Background and Preliminaries

Graphs are mathematical structures used to model pairwise relations between objects. In this structure, objects are represented as nodes that are connected by edges [120]. In order to preserve temporal information via graph representations, Kostakos [98] introduced the idea of temporal graphs. In this representation, contrary to static representations, the temporal relationship between nodes and the position of each node in the temporal context of the entire network is preserved. According to Kostakos [98], a temporal graph is constructed in three steps:

1. One temporal node per original node per point in time is created. Thus, node v_i in the original graph is represented as $v_{i,t}$ in the temporal graph where $t \in [1, .., T]$ are all the data points in which node v_i is communicated; e.g., Child-B in Figure 3.1 is represented by the set of instances $v_{B,t} = [v_{B1}, v_{B4}, v_{B6}]$.

2. Consecutive pairs $v_{i,t_x}, v_{i,t_{x+1}}$ with directed edges of weight $t_{x+1} - t_x$ for each set of instances v_i are linked, representing the temporal distance between the pair. For example, in Figure 3.1, the weight between nodes v_{B1} and v_{B4} is 3 time points.

3. The unweighted, undirected edges are used to link instantaneous communications between nodes. E.g., an interaction between Child-A and Child-B in Figure 3.1 at time t_1 is instantiated as an undirected link between Child-A and Child-B.

Therefore, each node v_i in Figure 3.1-b is converted to a directed chain of nodes $v_{i,t}$ that represent all temporal instances of nodes over time. Following this protocol, the temporal graph is generated as shown in Figure 3.1-c.

Furthermore, Kostakos [98] defined average geodesic proximity $G(v_i, v_j)$ on a temporal graph as the measure of "on average, how many hops is node v_i away from node v_j". This metric is formulated as follows:

$$G(v_i, v_j) = \frac{1}{n} \sum_{t=1}^{T} g(v_i, v_j, t), \tag{3.1}$$

where $g(v_i, v_j, t)$ is the shortest path, starting from node v_i at time t to the most accessible version of node v_j over time. $t \in [1, ..., T]$ is the set of time points that node v_i is active (or communicating). n is the number of finite $g(v_i, v_j, t)$ values, e.g., $g(A, D, t = 1) = 2 \;\; (i.e., v_{A1}, v_{A2}, v_{D2})$ in our example in Figure 3.1. In this metric, the lowest number of hops defines the most accessible instance.

Consequently, the $G_{in}(v_i)$ and $G_{out}(v_i)$ are calculated for node v_i as formulated in Equation 3.2:

$$\begin{aligned} G_{in}(v_i) &= \frac{1}{n} \sum_{j=1, j \neq i}^{M} G(v_j, v_i), \\ G_{out}(v_i) &= \frac{1}{n} \sum_{j=1, j \neq i}^{M} G(v_i, v_j) \end{aligned} \tag{3.2}$$

Where $G_{in}(v_i)$ and $G_{out}(v_i)$ are measures of "on average, in how many hops is v_i reached by the rest of the network" and "on average, in how many hops does v_i reach the rest of the network", respectively. n is the number of finite $G(v_j, v_i)$ and $G(v_i, v_j)$ values, and $v_j \in \mathbf{V} = \{v_1, v_2, ..., v_M\}, v_j \neq v_i$ is sets of nodes in a network with M nodes.

Similarly, average temporal proximity $P(v_i, v_j)$ is the measure of, "on average, the time it takes to go from v_i to v_j" and is formulated as follows:

$$P(v_i, v_j) = \frac{1}{n} \sum_{t=1}^{T} p(v_i, v_j, w_t, t), \tag{3.3}$$

where $p(v_i, v_j, w_t, t)$ is the weighted shortest path, starting from node v_i at time t to the most accessible version of node v_j over time. $t \in [1, ..., T]$ is the set of time points that node v_i is active (or communicated). n is the number of finite $p(v_i, v_j, w_t, t)$ values e.g., $p(A, D, w_t, t_1) = 1 \;\; (i.e., v_{A1}, v_{A2}, v_{D2})$ in Figure 3.1. The weight w_t equals the time difference between nodes at a given path. Consequently, the $P_{in}(v_i)$ and $P_{out}(v_i)$ are calculated as formulated in Equation 3.4:

$$\begin{aligned} P_{in}(v_i) &= \frac{1}{n} \sum_{j=1, j \neq i}^{M} P(v_j, v_i), \\ P_{out}(v_i) &= \frac{1}{n} \sum_{j=1, j \neq i}^{M} P(v_i, v_j) \end{aligned} \tag{3.4}$$

The P_{in} and P_{out} are measures of "how quickly, on average, v_i is reached by the rest of the network" and "how quickly, on average, v_i reaches the rest of the network". n is the number of finite $P(v_j, v_i)$ and $P(v_i, v_j)$ values, and $v_j \in \mathbf{V} = \{v_1, v_2, ..., v_M\}, v_j \neq v_i$ is sets of nodes in a network with M nodes. This metric estimates nodes' accessibility by obtaining its availability over time.

To better understand these two temporal metrics, we calculated them in the playground example presented in the Introduction (See Figure 3.1). Table 3.1 and Table 3.2 show the result of average geodesic proximity and average temporal proximity, respectively. The results demonstrate that, despite the static measures, Child-A and Child-B have different in and out measures (both geodesic and temporal). Child-A has the lowest G_{in} and P_{in} as all peers find the child shortly at the start of the break, and the highest G_{out} due to its absence in communication during the second half of the break. This might be because the child is more involved in solitary games, rather than being with peers during recess. The lowest values belong to Child-B in both G_{out} and P_{out} because Child-B accessed peers across the break and interacted with all peers.

Table 3.1: The result of average geodesic proximity on playground example

	G				G_{out}
	A	B	C	D	
A	0	2.33	1.33	2.67	2.111
B	1	0	1	2	1.333
C	1	1.33	0	3.33	1.889
D	1	1.5	3	0	1.833
G_{in}	1	1.722	1.778	2.667	

Table 3.2: The result of average temporal proximity on playground example

	P				P_{out}
	A	B	C	D	
A	0	0.33	1.67	0.39	0.295
B	0.14	0	0.16	0.14	0.148
C	0.15	0.18	0	0.32	0.219
D	0.32	0.14	0.52	0	0.325
P_{in}	0.202	0.218	0.281	0.284	

3.4.2 Spatio-temporal Proximity

Both of the metrics that are explained in the previous section overlook spatial dependency among nodes, and consequently, ignore the spatial behavior of individuals (e.g., whether a child is interacting in different areas over the course of recess, or the child stays in a single spot and interacts with peers who come over to that same spot). Obtaining such detailed spatio-temporal behavior is not possible via the existing measures (e.g., average geodesic proximity and average temporal proximity). To address this gap, we introduce spatio-temporal graphs and define average spatio-temporal proximity as $S(v_i, v_j)$, which takes into account spatial dependencies among nodes. By considering the spatial distances, e.g., Euclidean distance, the spatio-temporal proximity examines how spatially close the user is to their network. To create a spatio-temporal graph, we followed the same steps as in the temporal graph protocol, except in step (2), where the consecutive pairs $v_{i,s_{t_x}}, v_{i,s_{t_{x+1}}}$ are linked with the directed edges of weight $d(v_{i,s_{t_x}}, v_{i,s_{t_{x+1}}})$, which is the spatial distance between instances $v_{i,s_{t_x}}$ and $v_{i,s_{t_{x+1}}}$ (instead of temporal differences).

Accordingly, average spatio-temporal proximity is calculated based on this spatio-temporal graph, and allows the weights to be the spatial distances among the edges. This metric investigates "on average, how long (spatio-temporally) it takes for the network to reach v_i (S_{in})", and "on average, how long (spatio-temporally) it takes for v_i to reach the rest of the network (S_{out})". The average spatio-temporal proximity is formulated as follows:

$$S(v_i, v_j) = \frac{1}{n} \sum_{t=1}^{T} s(v_i, v_j, w_s, t), \tag{3.5}$$

where $s(v_i, v_j, w_s, t)$ is the weighted shortest path, starting from node v_i at time t to the most accessible version of node v_j over time. $t \in [1, ..., T]$ is the set of time points that node v_i is active (or communicated). n is the number of finite $s(v_i, v_j, w_s, t)$ values. The weight w_s is equivalent to the spatial distance between nodes v_i and v_j at a given path, assuming that the location coordinates of all the contacts are available.

Furthermore, the spatio-temporal proximity is calculated per pair of nodes in a community, and the in-degree (S_{in}) and out-degree (S_{in}) variants are calculated by taking the average over the in-reached and out-reached variables, respectively.

$$\begin{aligned} S_{in}(v_i) &= \frac{1}{n} \sum_{j=1, j\neq i}^{M} S(v_j, v_i), \\ S_{out}(v_i) &= \frac{1}{n} \sum_{j=1, j\neq i}^{M} S(v_i, v_j) \end{aligned} \tag{3.6}$$

The $S_{in}(v_i)$ and $S_{out}(v_i)$ are measures of "how long, on average, v_i is accessible by the rest of the network" and "how long, on average, it takes on average for v_i

to access the rest of the network". n is the number of finite $S(v_j, v_i)$ and $S(v_i, v_j)$ values, and $v_j \in \mathbf{V} = \{v_1, v_2, ..., v_M\}, v_j \neq v_i$ is sets of nodes in a network with M nodes. Similarly, the result of this analysis for the playground example is shown in Table 3.3. To calculate this table, it is assumed that the sandpit, bench, and entrance area are located at $(3, 1)$, $(2, 2)$, and $(1, 1)$ Cartesian coordinates, respectively, and that the interactions happened around these areas. As described

Table 3.3: The result of average spatio-temporal proximity on playground example

	S				S_{out}
	A	B	C	D	
A	0	0.85	0.1	0.83	0.594
B	0.141	0	0.1	0.48	0.241
C	0.1	0.43	0	1.2	0.579
D	0.1	0.7	0.2	0	0.333
S_{in}	0.114	0.66	0.133	0.84	

in Table 3.3, Child-A has the lowest S_{in}: during time points when the child was available (i.e., T1, T2, and T3), all other peers were in the same spot (sandpit) and could access the child with a short distance to travel. Meanwhile, Child-D got the highest S_{in} because the child was only available at T2 and T6, and if anyone wanted to reach Child-D after T2, they had to pass a long spatio-temporal route (i.e., temporally from T2 to T6, and spatially to the entrance, in order to find the child at the Entrance on T6). Regarding the S_{out}, Child-A has the highest score because some peers were not presented at T2 and T3 in the sandpit. Therefore, the child had to wait and reach them in different locations. In contrast, Child-B could more easily access peers (lowest S_{out}) because the child was present across recess time in different locations. Thus, our proposed metric quantifies accessibility by unifying the spatial and temporal dimension of interactions, and provides unique knowledge on how individuals and their networks utilize the spatial environment to make contacts over time.

3.5 Experiments

We investigated the performance of the proposed metric in identifying social groups, based on the most accessible peers. Specifically, we were interested in knowing whether adding spatial context to the existing temporal metrics (i.e., geodesic and temporal proximity) could help better identify groups in a network.

We evaluated the proposed methodology by utilizing a clustering performance task. Here, we first explain our experimental design and introduce the datasets adopted in our experiments. Then, the evaluation metrics are presented. Next, we show the result of our analysis in two sections: The first part evaluates the performance of a clustering task using the proposed metric, compared with the existing temporal metrics as the baseline. The second part analyses children's behavior in the schoolyard dataset.

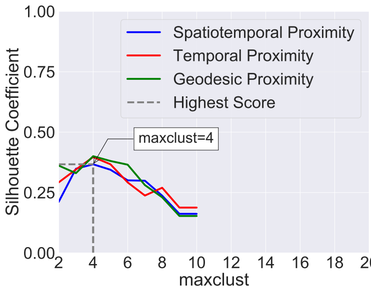
3.5.1 Experiment Settings

Experiments were designed using the Python programming language. The code has been made available in a public repository[1].

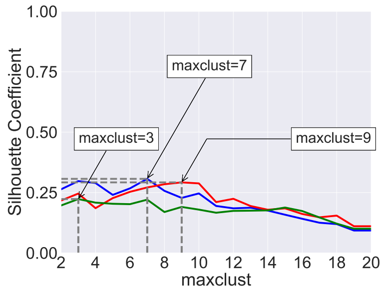
The NetworkX module was used to execute social network analysis, and Dijkstra shortest-path algorithm [121] was used for the calculation of proximity metrics (i.e., geodesic, temporal, and spatio-temporal proximity) on the spatio-temporal graph. The SciPy and Scikit Learn packages were used to perform clustering tasks and evaluate their performance, respectively. Hierarchical clustering [122] was used to cluster the proximity matrices, which was a natural choice for clustering data in the form of an adjacency matrix, especially when the dataset size is small. In addition, hierarchical clustering serves best in exploratory analysis where there is no specific outcome variable defined. For this purpose, first, the Pearson correlation coefficients of the pairwise matrices were calculated separately across in-degree and out-degree components to acquire how in-degree proximity (i.e., accessing the child by the network), and out-degree proximity (i.e., accessing the network by the child), were correlated among children. Consequently, we obtained a symmetric correlation coefficient matrix (CCM) per in-degree and out-degree metrics, thereby determining the strength and direction of the pairwise relationships among children. Next, the upper triangular area of the in-degree CCM was merged with the lower triangular area of the out-degree CCM, to create a non-symmetric CCM that described both in-degree and out-degree correlations among children. Such a non-symmetric CCM was used in the hierarchical clustering algorithm to identify groups based on mostly correlational relationships. Furthermore, the obtained matrix was used in creating heatmaps to visualize the relationships between children in a more organized way, which aids in data exploration and analysis.

In designing the hierarchical clustering algorithm, three different linkage methods, including complete, average, and average group, were investigated to understand how different strategies in creating the linkage matrix impact the performance

[1] https://github.com/maaeedee/spatio-temporal-network.git



(a)



(b)

Figure 3.2: The maximum Silhouette coefficient score is used to find the optimum 'maxclust' value t for the hierarchical clustering algorithm on the range of $t \in [2, N]$, in which N is the total number of data points for (a) PG-1 ($N = 10$) and (b) PG-2 ($N = 20$). Accordingly, the 'maxclust' is selected as 4 for all three measures in PG-1 and as $G = 9, P = 3, S = 7$ for PG-2 based on the maximum Silhouette coefficient score.

of the clustering algorithm. In short, the 'complete' linkage method defines the distance between two groups as the distance between the most distant pair of points, while the 'average' defines this distance as the average pairwise distance between data points in two groups, and 'average group' defines it as the average pairwise distance between all data points when two groups are merged. Moreover, the correlation distance was used to ensure that the clustering algorithm considered the strength and direction of the correlations between variables. The 'maxclust' was selected as the criterion, which finds a minimum threshold to ensure the cophenetic distance between any two original points in the same flat cluster is no greater than the threshold, and no greater than t flat clusters are formed. Since the number of clusters was unknown, the optimum value of t was chosen based on the result of clustering performance, using the Silhouette coefficient score [123] on the range of $t \in [2, N]$, in which N is the total number of data points. This range was chosen based on the assumption that there were at least two clusters in a playgroup, and that, in the worst-case scenario, there was no child in the same cluster with peers. Since 'complete' linkage, on average, performed slightly better than the other two linkage methods (as shown in Table 3.5), this linkage matrix was used to obtain the optimum number of clusters. Figure 3.2 shows results for both playgroups. Based on these figures, the 'maxclust' was selected as 4 for all proximity measures in PG-1 and $[G = 9, P = 3, S = 7]$ for PG-2 based on the maximum Silhouette coefficient score.

3.5.2 Schoolyard Dataset

We collected spatio-temporal data via wearable sensors from 32 children, aged between 4 and 12 years old, in two playgroups, PG-1 and PG-2, during recess in the schoolyard of a primary special education school in the Netherlands. These data are now publicly available, so other researchers may replicate our results[2].

Playgroups PG-1 and PG-2 were children from junior and senior grades, respectively, who were assigned to designated playgrounds. As shown in Figure 3.3, PG-1 used areas I, II, and III, which included a soccer field, bench, climbing structure, and sandpit. PG-2 spent recess time in areas IV and V, where a bench, table tennis, climbing frame, and soccer field were located. Before starting each break, children were asked to wear a belt with a mounted proximity tag and a GPS logger. The proximity tags detected face-to-face contacts, and the GPS loggers recorded the geographic locations of the users. Specifications for the datasets are described in Table 3.4. In addition to sensor data, video recordings, and field observation were conducted simultaneously to gain insight into the psychological aspects of children's interactions and social behavior. Parental informed consent was obtained before conducting the data collection. Approval for the study was obtained from the Leiden University Ethical Committee.

Table 3.4: Specification of Schoolyard Datasets for playgroups PG-1 and PG-2

Datasets	#Groups	#SessionsxDuration	#Participants	Age (M $\pm$ SD)
PG-1	2	2 x 30min, 1 x 15min	11	5.7 $\pm$ 0.96
PG-2	3	2 x 15min	22	10.43 $\pm$ 0.77

In order to evaluate the effectiveness of our proposed method and verify the results, we need ground-truth data (i.e., unbiased and true values serving as a reference for evaluating the method). The ground-truth informs us about the true pairwise relationship among children, e.g., the true values on how often a child was in meaningful interaction with peers, where precisely in the schoolyard, and for how long. This data could be presented, for example, in the form of pairwise group membership over time by defining 'group' as peers who are most accessible by their group mates. Due to privacy concerns regarding any child population, existing videos associated with the data collection were recorded only from focal areas, following the approved ethics application. This made it practically impossible to create ground-truth for the collected sensor data and to systematically verify the

[2] https://doi.org/10.34894/ZCPXDW

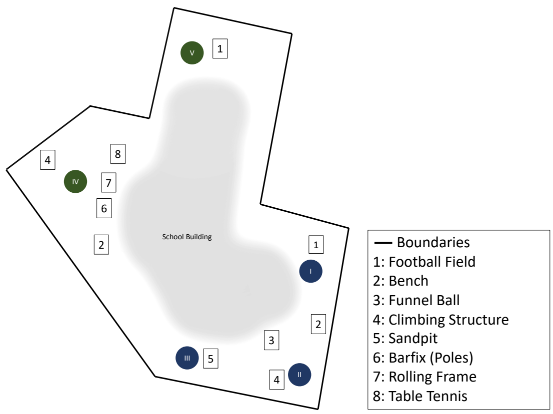


Figure 3.3: Schoolyard floor plan: Junior grades spend their recess time in areas I, II, and III, while senior grades are in areas V and IV. The schoolyard environment has different play structures, as described in the legend, which children may use during recess.

results. Moreover, evaluating the concept of pairwise proximity or accessibility over time might not be easily understood and coded by monitoring video recordings or direct observations, especially with regard to children's age groups who often tend to change their peer group and attend different types of play over the course of recess time. Yet, the collected video recordings and the field observation reports allowed us to discover certain behaviors (e.g., identifying the most visited areas, popular activities (or games) among children of a certain group, etc.), to draw a general picture of schoolyards activities, and to verify the obtained results to some extent. Therefore, the experiments in this section are at an exploratory level considering different age groups and playground designs.

3.5.2.1 Data Preparation and Pre-processing

Proximity Data: Wearable proximity tags were used to detect face-to-face contact between children during recess. If any pairs of proximity tags (i.e., children) were at a distance of up to $1.5ms$ in the degree of about $60-70°$, then they registered each other's unique code via Bluetooth. They wirelessly sent this detection to two base stations installed in two different locations in the schoolyard. The installation locations were estimated beforehand to obtain the highest coverage over the schoolyard by considering the fact that each base station covers an area of up to $25m^2$.

Hence, some children might have a low detection rate for various reasons, e.g., playing far away from the base stations or tag malfunctioning. To compensate for this error, the detection rate is calculated per child by dividing the duration of detection by beagle over the total duration of the break. Only children with a detection rate higher than 50% are included for further analysis. As a result, three children from PG-2 who had a low detection rate were excluded from further analysis. The data reported in Table 3.4 is after this exclusion. The obtained contact data is then fused with GPS data to obtain the location of interactions

GPS Locations: The GPS logger(i-gotU GT-120 USB) recorded location coordinates in Longitude and Latitude per second. The GPS locations were converted afterward to Cartesian space and merged with proximity contacts to obtain the location of interactions. As suggested in the literature [91, 124, 125], we adopted Euclidean distance to examine the spatial distances between temporal nodes (w_s in Equation 3.5). In our study, it was assumed that children in face-to-face contact are almost in the same location, to use the data more efficiently. Therefore, the available GPS data of one of the pairs in interaction was used for the other, in case the GPS data were missed. In addition, a 10-second window was used to search for the location of individuals in an interaction, and the median values were used to report the location coordinates at the given time. This merged data was used in our analysis to create the spatio-temporal graph.

3.5.3 Evaluation Metrics

This section describes the evaluation metrics used to examine the performance of a clustering task. Since no ground-truth data were available for the schoolyard dataset, we decided to use the following evaluation metrics, which do not require ground-truth but still provide a quantitative assessment of the quality of the clusters found in these datasets when different proximity metrics were used. We expect to observe clusters with higher quality when more informative proximity metrics are used.

1. **Silhouette score** [123] indicates how well clusters are separated from each other. This metric is defined as follows:

$$SC = \frac{1}{N}\sum_{i=1}^{N}\frac{(b_i - a_i)}{max(a_i, b_i)}$$

   where a is the mean distance between sample i and all other points in the same cluster, and b is the mean distance between sample i and all other points in the next nearest cluster. N is the total number of points.

2. **Calinski-Harabasz score** [126] is the ratio between the within-cluster dispersion and the between-cluster dispersion, and is defined as follows:

$$s = \frac{\mathrm{tr}(B_k)}{\mathrm{tr}(W_k)} \times \frac{n_E - k}{k - 1}$$

where n_E and k are the data and cluster sizes, respectively. $\mathrm{tr}(B_k)$ is the trace of the between-group dispersion matrix, and $\mathrm{tr}(W_k)$ is the trace of the within-cluster dispersion matrix defined as follows:

$$W_k = \sum_{q=1}^{k} \sum_{x \in C_q} (x - c_q)(x - c_q)^T, \quad B_k = \sum_{q=1}^{k} n_q (c_q - c_E)(c_q - c_E)^T$$

where C_q, is the set of points in cluster q, with cluster center equal to c_q. c_E is the centroid of cluster E and n_q is number of points in cluster q.

3. **Davies-Bouldin score** [127] is defined as the ratio between the cluster scatter and the cluster's separation. A lower value of this metric will indicate a better clustering and it can be calculated as follows:

$$DB = \frac{1}{k} \sum_{i=1}^{k} \max_{i \neq j} R_{ij}, \quad R_{ij} = \frac{s_i + s_j}{d_{ij}}$$

where s_i is the diameter of cluster i and d_{ij} is the distance between cluster centroids i and j. DB is the Davis-Bouldin index.

3.5.4 Results and Discussion

In this section, we report the results of our experiments in two main areas: evaluation of the performance of the proposed metric by obtaining the quality of clusters in a hierarchical clustering task (Clustering Performance Evaluation), and the in-depth analysis of our Schoolyard Datasets, to see how children interacted with peers and utilized the physical environment.

3.5.4.1 Clustering Performance Evaluation

Table 3.5 shows the results of the clustering performance task using Silhouette, Calinski–Harabasz, and Davies-Bouldin scores in clustering CCM of pairwise relation. In pairwise relations of PG-1, the geodesic proximity and temporal proximity obtained slightly higher Silhouette and Calinski–Harabasz scores, and lower

Davies–Bouldin scores than spatio-temporal proximity in the three linkage matrices. The superiority of geodesic and temporal proximity might be due to the fact that most children in PG-1 played with bikes across the playground, according to the field observations during the data collection. This activity uniformly increased face-to-face contact with all other children across the playground, thereby making them uniformly accessible. Therefore, pairwise spatial proximity might not be influential in describing pairwise proximity for PG-1. Meanwhile, in the senior group, PG-2, our proposed spatio-temporal measure and temporal measure identified clusters with higher quality than geodesic proximity by obtaining higher Silhouette and Calinski-Harabasz scores. Whilst the result was not consistent for Davies-Bouldin scores. This competitive result in temporal and spatio-temporal metrics indicates that in addition to focusing on the temporal context, including the spatio-temporal contexts of interactions contributed to describing children's social behavior and their position in the social network in PG-2. The availability of a larger playground area, the tendency to play in group settings in certain areas in senior groups, and more interest in stationary activities such as sitting and chatting with peers might be the reasons why the spatial context provided a positive impact on the performance of the clustering algorithm of this playgroup, compared with PG-1.

As Table 3.5 shows, the performance of linkage methods was inconsistent across different metrics and evaluation scores. It is important to note that the performance of different linkage methods can vary depending on the specific dataset and the nature of the underlying data. Therefore, to optimize clustering performance, the linkage method could be defined as a hyperparameter and optimized specifically per dataset. In the present study, we adopted the result of the 'complete' linkage matrix to conduct the analysis on the schoolyard dataset in the next sections of the analysis.

Overall, our results highlight the value of including the spatial context in examining individuals' social behavior, and in assessing their position in the network.

3.5.4.2 In-Depth analysis of Schoolyard Data

In this analysis, accessibility is defined in three dimensions: (1) *geodesic proximity* - $G(v_i, v_j)$, (2) *temporal proximity* - $P(v_i, v_j)$, and (3) *spatio-temporal proximity* - $S(v_i, v_j)$. Each dimension provides unique knowledge on how individuals and their networks were accessible in the geodesic space, over time, and in the spatio-temporal context. Therefore, "accessibility" will heretofore refer to all three of these dimensions, throughout the rest of this paper.

As previously discussed, the CCM was used to create heatmaps as depicted in Figure 3.4 for both playgroups. The x and y axis show children from different groups

Table 3.5: The result of clustering performance in pairwise relations, based on CCM of geodesic proximity $G(v_i, v_j)$, temporal proximity $P(v_i, v_j)$, and spatio-temporal proximity $S(v_i, v_j)$ for three different linkage methods: complete, average and average group. The best performance per linkage method is indicated in boldface, and the best performance among all linkage methods is shown with $(*)$. a-group stands for Average group.

Data	**Metric**	**Silhouette**			**Calinski-Harabasz**			**Davies-Bouldin**		
		complete	**average**	**a-group**	**complete**	**average**	**a-group**	**complete**	**average**	**a-group**
PG-1	$G(v_i, v_j)$	**0.400***	0.380	0.312	7.865	**9.763***	6.397	0.722	**0.622***	0.994
	$P(v_i, v_j)$	**0.397**	**0.397**	**0.397**	**9.166**	**9.166**	**9.166**	**0.734**	**0.734**	**0.734**
	$S(v_i, v_j)$	**0.367**	**0.367**	0.313	**7.363**	**7.363**	6.369	**0.814**	**0.814**	1.170
PG-2	$G(v_i, v_j)$	0.223	**0.241**	0.212	**5.879**	5.507	5.324	1.454	**0.893**	0.977
	$P(v_i, v_j)$	**0.292**	**0.292**	0.249	**8.276***	**8.276***	8.006	**0.666**	**0.666**	1.465
	$S(v_i, v_j)$	0.307	**0.308***	0.263	**7.438**	6.713	6.637	**0.608***	0.733	1.583

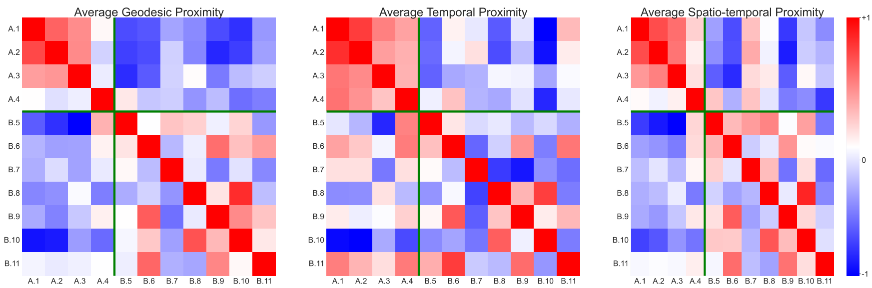
in this figure. Specifically, there are two groups (A and B) in PG-1 and three groups (A, B and C) that have similar break schedules, separated by green lines. The first digit of each child's identifier indicates their group. The results are shown as a heatmap based on the correlation coefficients matrix in all three measures (i.e., average geodesic proximity, average temporal proximity, and average spatio-temporal proximity). The grids are colored based on obtained correlation coefficients, such that blue shows negative correlations, red shows positive correlations, and white shows no correlation among pairs of children, following the color bar.

We further quantified and visualized the results of these heatmaps using the box plots to better understand the relations in each playgroup. Figure 3.5 shows the box plot of inter-group and cross-group correlation of the results per proximity measure in both playgroups.

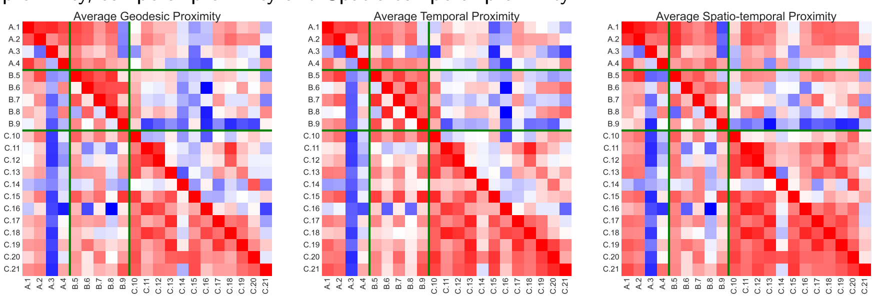
In Figure 3.5-a, we observed that in PG-1-A, the highest correlation was found in pairwise temporal proximity, compared with the other two proximity measures. Meanwhile, in group PG-1-B, spatio-temporal proximity scored the highest correlation, indicating that spatial context could better describe the accessibility among children of this group. In addition, the average correlation is higher in PG-1-A compared with PG-1-B, among all proximity measures. This could be because of the smaller group size in PG-1-A. Overall, there is a positive correlation between the proximity measures of children from the same group and a negative correlation among children from different groups, except in the temporal proximity of PG-1-B, where the average cross-group correlation is higher than the inter-group correlation in the boxplot. This is also illustrated in Figure 3.4-a, where several negative correlations (cells in blue) were observed among children in PG-1-B. Nonetheless, no pronounced patterns between in-degree and out-degree proximity were found in this playgroup.

Figure 3.5-b shows that in PG-2 the positive and negative correlations are more uniformly distributed across inter-groups and cross-groups pairs. Yet, in all three groups (A, B, and C) and across all proximity measures, the average inter-group correlation is higher than the cross-group measures. This means that, in general, children are more accessible by peers from the same group, and it is easier to access peers from the same group in all three dimensions of geodesic, temporal, and spatio-temporal proximity. Yet, some children appear as outliers, not following the general pattern in the heatmap visualizations (e.g., A4 and B11 in PG-1, A3 and C14 in PG-2 as illustrated in Figure 3.4).

Furthermore, the triangular patterns are more pronounced in PG-2 than in PG-1. This shows a considerable difference between proximity towards a child from the peer group and proximity towards peers by a child. Specifically, in all three measures, children in PG-2-A and PG-2-B have a positive correlation in their in-degree components (i.e., more red cells in their upper triangular heatmap). In comparison, in PG-2-C the positive correlation is observed in out-degree measures (more red cells in their lower triangular heatmap). This means that the peer networks in groups A and B had quicker access to a child, compared with the child accessing the network. On the contrary, in group C, individuals had quicker access to peer groups.



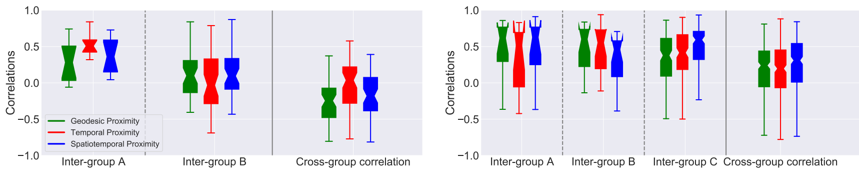
(a) PG-1 (groups A and B): the heatmap of pairwise accessibility in Geodesic proximity, temporal proximity and Spatio-temporal proximity.



(b) PG-2 (groups A, B, and C): the heatmap of pairwise accessibility in Geodesic proximity, temporal proximity and Spatio-temporal proximity.

Figure 3.4: The heatmap of Schoolyard Dataset for two playgroups: comparing the correlation coefficients matrix of geodesic proximity, temporal proximity, and spatio-temporal proximity among children in playgroup PG-1 and PG-2 (x.code = group id (A, B, or C). Child identifier code). The playgroup consists of several groups (A, B, and C) divided by green lines.

Furthermore, to better understand the impact of physical space in shaping identified clusters, we extracted the GPS location of children per cluster across one of the school breaks, and visualized the kernel density estimation (KDE) of the GPS locations. KDE represents the distribution of GPS locations using a continuous



(a) (b)

Figure 3.5: The correlation of pairwise proximity for two playgroups: comparing the CCM of geodesic proximity (in green), temporal proximity (in red), and spatio-temporal proximity (in blue) among children in playgroup PG-1 and PG-2. On average, children in both groups have a higher correlation with peers from the same group and a lower correlation with peers from different groups. This difference is more significant in PG-1-A and less significant in senior playgroup, i.e., PG-2.

probability density curve in two dimensions. The result of this analysis is depicted in Figure 3.6 and Figure 3.7 for PG-1 and PG-2, respectively.

As shown in Figure 3.6, we identified four clusters via hierarchical clustering of pairwise spatio-temporal proximity. Each cluster shows how the most accessible peers within each cluster utilized the playground. As a result, in *Cluster 1*, children were most accessible in the communal area where most play structures were quickly reachable (e.g., climbing frame, funnel ball, sandpit, and bench). Children in *Cluster 2* and *Cluster 4* often played in sandpits and quickly accessed their peers playing in the same location. According to field observations, children in PG-1 often attended to wheeled movements. This is observed in *Cluster 1* and *Cluster 3*, in which children widely used the playground via wheeled mobility toys (e.g., bikes, steppers, etc.). This finding shows how the available play structures impacted children's use of space, and could determine group activities in a network.

The impact of available play structures was even more evident in PG-2. In Figure 3.7, hierarchical clustering of pairwise spatio-temporal proximity identified seven clusters. In Cluster 2, with the most members, children often had face-to-face contact with peers around the play structures. This could, for instance, be due to using several play structures during the break (e.g., climbing frame, bench, and bar fixes) or being involved in an active game that required movements. Thus, the highly supplied area by the play structures attracted more children, and more conscious or unconscious contacts happened in this cluster. *Cluster 1* and *Cluster 3* belong to children who stayed around and in the climbing frame to climb, chat, or to simply observe other peers. The use of space as influenced by a specific play

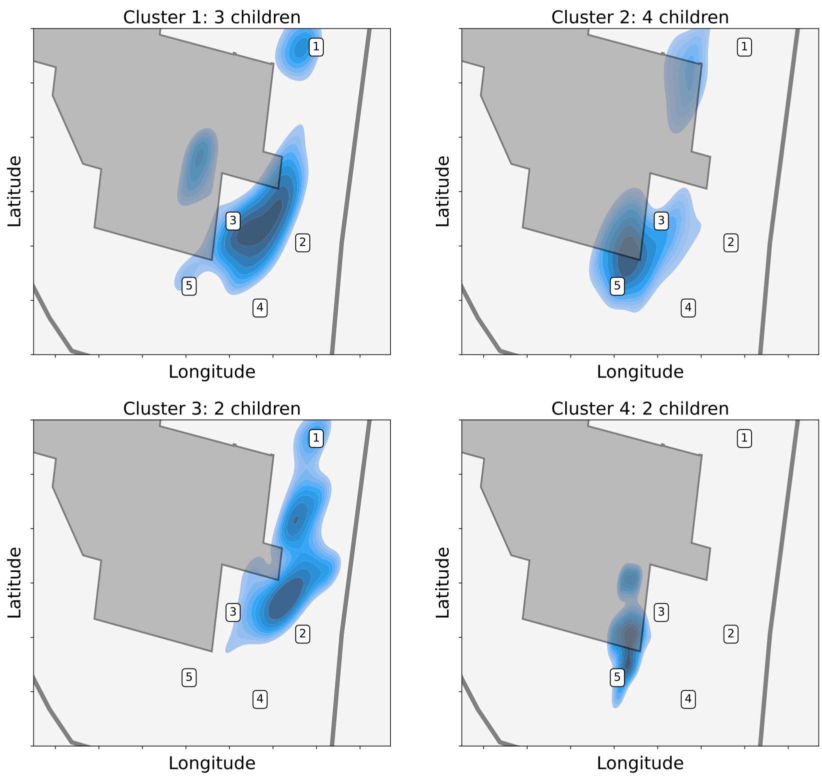


Figure 3.6: The GPS location of children in PG-1 is mapped to the school floor plan. The KDE plot in blue shows the spatial density of the GPS locations in the identified clusters, following the color bar.

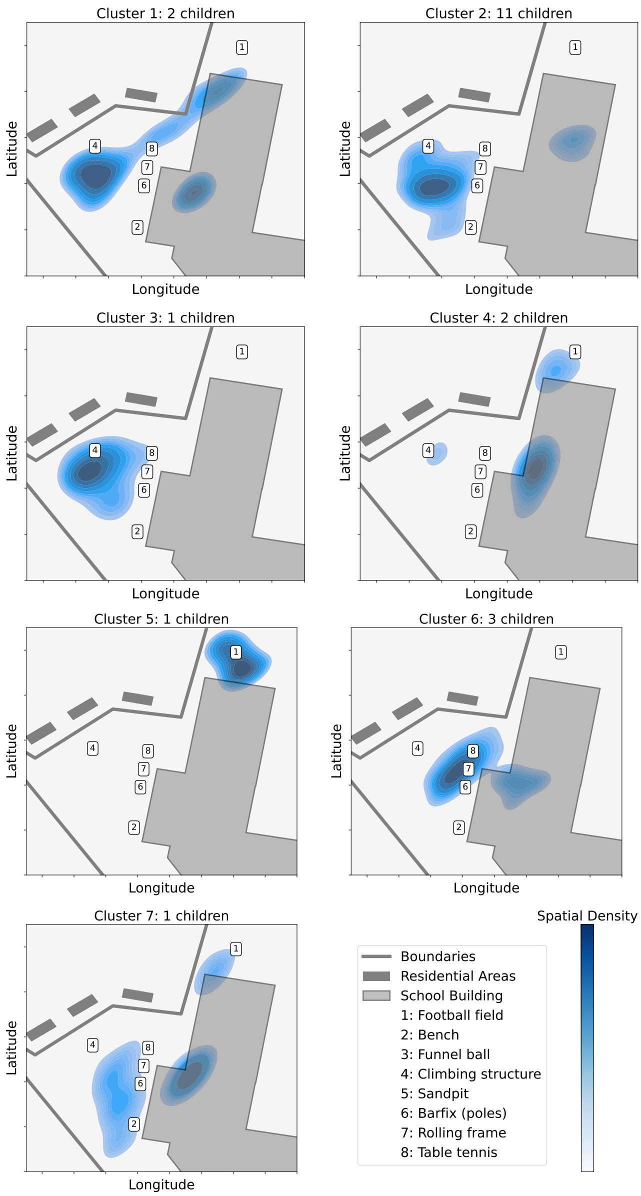


Figure 3.7: The GPS location of children in PG-2 is mapped to the school floor plan. The KDE plot in blue shows the spatial density of the GPS locations in the identified clusters, following the color bar.

structure was also observed in *Cluster 5* and *Cluster 6*, where the football field and the attached structures (i.e., table tennis, rolling game, and bar fix) were located. A higher number of clusters in this playgroup, compared with PG-1, occurred due to a higher number of participants in PG-2; more extensive physical space provided to this playgroup during recess; and differences in their age group, as senior grades tend to have more group-oriented activities, such as talking on a bench or table tennis, compared with junior grades.

3.6 Conclusions

The present study features a novel spatio-temporal proximity metric, which retains spatial information about the underlying temporal dynamics. The main strength of this metric, compared to the previously proposed temporal proximity metrics, lies in the ability to account for the role of the physical environment when analyzing an individual's temporal behavior. We applied the proposed metric to a case study, in an effort to understand children's social behavior in a schoolyard environment.

We evaluated the quality of the proposed metric by performing a downstream clustering task. Our results show that the proposed spatio-temporal metric uniquely quantified an individual's relationships over time, concerning the spatial context of their activity. This led to better clustering performance when compared to geodesic proximity metrics that were applied to senior groups in our study. In junior groups, the spatial context did not positively impact clustering performance.

We also introduced two new datasets collected from children playing in schoolyards during recess, using wearable proximity tags and GPS loggers. We made these datasets available for other researchers to study social dynamics. In this case study, we first examined pairwise accessibility based on the calculated proximity values among children in a playgroup. Our results showed that, in general, children were more accessible to peers from the same group. Similarly, children more easily accessed peers from the same group. Second, we analyzed the impact of the physical environment and its characteristics on forming the most accessible clusters. Our findings show that in both of the playgroups we studied, the most accessible clusters utilized common areas in the playground, which were often located around one or several play structures that the schoolyard featured.

In summary, we found that application of a novel spatio-temporal proximity metric to data collected via wearable proximity sensors offered valuable insight into the spatio-temporal accessibility of children in schoolyards. The proposed method could be used to develop and evaluate targeted interventions aimed at creating a more accessible and inclusive environment, for example, by designing new play structures or interactive games aimed at increasing pairwise accessibility among all children. Moreover, the proposed method is applicable beyond schoolyards and can be used in diverse scenarios, including analyzing employee behavior in an office setting, tracking athletes' movements on sports fields, and monitoring the well-being of elderly individuals in a nursing home. The present study used the proposed metric in the context of a case study at an exploratory level. Yet, making broad generalizations about children's behavior requires a larger sample of data. Collecting a larger sample was not possible due to the COVID-19 crisis, which resulted in school closure and implementing strict protocols during school re-openings, but it currently remains one of the major focuses of the research team.

One of the limitations of this research is the absence of ground-truth data due to privacy concerns and the complexity of defining "truth" for the examined variables. Specifically, in order to provide ground-truth for children's spatio-temporal proximity, we need to know the details of their group behavior in time and space, e.g., the true values of children's pairwise meaningful interactions, the frequency of contacts and their accurate location in the schoolyard. Thus, it is important to consider some levels of uncertainty and the potential for alternative interpretations, e.g., uncertainties in sensors operations might conclude a different picture of schoolyard activities or use of some specific areas might be due to the impact of external events such as weather conditions, etc. Another limitation in the proposed method is the time complexity of shortest-path algorithms, which is $\mathcal{O}(V \log V + E)$ for the Dijkstra shortest-path algorithm, where E and V denote the number of edges and the number of nodes, respectively. Thus, it is computationally expensive to estimate the proximity metrics when the data is gathered from larger samples (higher number of users) and over a longer period of time (higher number of temporal edges). Employing shortest-path estimation algorithms and parallel computing is recommended to enhance the scalability of the proposed method for larger graphs in future research. Additional research is needed in order to understand factors contributing to an inclusive environment, to investigate how the proposed metrics describe such an environment, and to prioritize interventions using the proposed metrics in a recommender system. Moreover, future research might explore the Kinetic data structures, a data structure to track an attribute of a moving geometric system, to capture the spatio-temporal changes of moving users.

CHAPTER 4

Social Connectedness and Loneliness in School for Autistic and Non-autistic Children

The contents of this chapter are based on the following accepted study (in press):

- Y.-T. Tsou, M. Nasri, B. Li, E. M. Blijd-Hoogewys, M. Baratchi, A. Koutamanis, and C. Rieffe, "Social connectedness and loneliness in school for autistic and allistic children," *Autism*, p. 13623613241259932, 2024.

  DOI: 10.1177/13623613241259932.

Abstract

Autistic children are often reported less socially connected, whilst recent studies show autistic children experiencing more loneliness in school than allistic (i.e., non-autistic) children, contradicting the traditional view that autistic children lack social motivation. This study aimed to understand individual differences in how social connectedness is construed between and within groups of autistic and allistic pupils using a multi-method approach. Forty-seven autistic and 52 allistic classmates from two special primary schools participated (8-13 years). Proximity sensors worn by pupils on playgrounds during recess measured (i) total time in face-to-face contacts, (ii) number of contact partners, and (iii) centrality in playground networks. Peer reports measured (iv) reciprocal friendships and (v) centrality in classmate networks. To evaluate their feelings of connectedness, pupils rated the level of loneliness in school. Compared to allistic pupils, autistic pupils had fewer reciprocal friendships, but similar total time in social contacts, number of partners, classmate/playground centrality, and levels of loneliness. Lower levels of loneliness related to higher classmate centrality in autistic children, but longer time in social contacts in allistic children. For these autistic children, being liked as a part of a peer group seems a key factor. Understanding relevant differences in children's needs could lead to a more welcoming school climate.

4.1 Introduction

Many previous studies reported that autistic children have fewer social connections. Yet, recent studies also show that autistic children more often feel lonely in school than allistic (i.e., non-autistic) children. This outcome seems to go against the traditional view that autistic children do not desire to have social connections. Therefore, this study aimed to find out how autistic and allistic children feel about their social connections. We included 47 autistic and 52 allistic children from two special education primary schools (aged 8-13 years). We tested their social connections and loneliness in school through a new approach. This new approach includes questionnaires and sensors for tracking social contacts on playgrounds during school breaks. We found that allistic children felt more loneliness when they spent little time in social contact during school breaks. Yet, autistic children felt more lonely when their peers did not like to play with them. For these autistic children, feelings of loneliness may go beyond face-to-face contact. Being liked as part of a peer group was key. Understanding differences in children's needs can lead to a more effective design for a welcoming school climate.

4.1.1 Social connectedness and loneliness in school for autistic and allistic children

School is the main setting where most children experience social interactions outside the family circle. For a child to learn and practice social skills, these social interactions are crucial. However, the nature of these interactions also relates to how socially connected a child feels at school [128, 129]. When there is a mismatch between the desired and actual amount of social connectedness – i.e., when a child is cognitively aware of the unmet desire in the quality and quantity of their social connections – feelings of loneliness arise [130]. Feelings of loneliness in school can be stressful and painful, and they can contribute to mental health problems, e.g., symptoms of depression and low self-esteem [131, 132]. Yet what is less well-known is the extent to which individual children may differ in how they construe social connectedness, especially among autistic children. For example, extensive literature has suggested that autistic pupils lack the desire to build social connections [133–135]. Yet other studies have found elevated levels of loneliness in school among autistic adolescents, compared to allistic (i.e., non-autistic) peers, thereby challenging that view [29, 136] also see [137, 138] for arguments against the "social motivation" view). These discrepancies reflect a gap in the literature regarding individual differences in how autistic and allistic pupils construe their own social connectedness.

In this study, we focused on groups of primary-school autistic and allistic pupils and distinguished two types of social connectedness: i) physical connectedness, i.e., pupils' physical proximity with their peers in school at recess, and ii) emotional connectedness, i.e., the peer connections with which pupils identify. Further, we examined how pupils felt about their social connectedness by measuring their subjective feelings of loneliness. To capture the social dynamics in the school environments, we uniquely employed a multi-method approach that combined self-report, peer nomination, and wearable sensor technology.

4.1.2 Social Connectedness of Autistic Pupils

Pupils may build social connections simply by being in proximity, such as playing next to a peer on the playground. Such opportunities for physical contact alone could suffice to promote peer interaction [139] and foster mutual understanding (see "contact theory" [140]). This type of physical connectedness in school could be particularly relevant for autistic pupils, as it may allow them to remain part of the group and learn social skills without becoming overwhelmed by the social demands required for building more intimate relationships. Despite this, current empirical evidence from observations shows that autistic pupils experience less physical contact in primary school compared to allistic peers. They more often spend time alone during recess, engaging in unoccupied or solitary activities, and initiate or respond to social interactions less frequently [141–144]. When a social connection becomes meaningful to pupils and is acknowledged [145, 146], a psychological bond and preferences may form among them, leading to a sense of emotional connectedness. Past research has shown that when asked to identify their social connections, autistic pupils tend to receive fewer nominations from peers as social group members, often occupying more peripheral positions in a peer network, and engage in fewer reciprocal friendships, compared to their allistic peers (e.g., [144, 147–150]. Additionally, they are less often perceived as "someone to hang out with" by allistic peers [148], and are more often considered as not preferred across their primary school years [150]. This pattern seems to primarily apply to autistic boys, who are more frequently rejected, while autistic girls are more often overlooked (i.e., not mentioned in any types of nominations [151]). These findings touch upon the idea that autistic pupils may feel less emotionally connected to their peers in school.

However, different results also surfaced when the perspectives of autistic pupils themselves were involved. Reports show that autistic pupils in primary school perceive themselves as socially involved because they nominate more friends and "buddies" than their allistic peers, although these nominations are more often unreciprocated [148]. At the same time, qualitative evidence indicated that many autistic pupils reported having one friend and being satisfied with the friendship [152]. For these autistic pupils, qualities such as shared interests, trust, and companionship seemed more important in their peer relationships, compared to the other qualities like reciprocity and closeness that were often valued by allistic pupils [152–155] also see [156] for a review). Learning from autistic pupils' varying experiences in school is thus crucial.

4.1.3 Loneliness in Autistic Pupils

While it appears that primary-school autistic children are lower in their physical and emotional connectedness in school than are allistic peers, the question is: are these children alone but satisfied with their level of social connection, or are they experiencing an unmet need and feeling lonely? To the best of our knowledge, only three studies directly compared levels of loneliness between autistic and allistic pupils in primary schools, via standardized self-report questionnaires, and no group differences were reported (aged 7-11 years; [136, 147, 148]). This is likely because autistic pupils at this age do see themselves as socially involved, as discussed above [148]. Moreover, feelings of loneliness were found to be unrelated to children's overall friendship quality or to how prominent they were in a peer group, both in autistic and allistic children [147, 148]. Nevertheless, in studies that also included adolescents, self-reported levels of loneliness in autistic participants were consistently higher than in their allistic peers [29,136,142,157–159]. In adolescence, most pupils experience a transition in their social environment, where peers become the primary partners in daily social interaction [160]. This transitional period can be particularly challenging for autistic adolescents, given the heightened expectations from the social environment. Notably, different factors seem to be relevant to these adolescents' reported loneliness. For allistic adolescents, higher levels of loneliness were related to fewer intimate and prosocial interactions with peers, while these relations were not observed in autistic adolescents [29, 142, 157]. Rather, autistic adolescents felt lonelier when their social networks did not provide a sense of togetherness and safety, i.e., when they experienced lower levels of trust and companionship in their friendships and more limited school participation [29,142]. Thus, when we want to gain a better understanding of the factors affecting loneliness in autistic youth, we need to consider possible individual differences and an approach that can capture dynamic features for social connectedness.

4.2 Present Study

Peer interaction is essential in most children's school life. But how do they feel about the connections that result from this interaction, and do differences such as having autism lead to different ways of viewing these connections? Such questions are important to answer, as they extend our knowledge on how children's social wellbeing in school may be enhanced.

In this study, we aimed to assess differences in how social connectedness was construed, both between and within groups of primary-school autistic and allistic pupils, respectively. We distinguished between physical and emotional social connectedness, and examined their relationships with pupils' feelings of loneliness, to understand the potential sources of unmet social connectedness needs. To this end, we recruited two special education schools, a setting where autistic children were with other neurodivergent peers in their class and where their needs are better addressed, compared to most mainstream schools investigated in many prior studies. Moreover, we adopted a multi-method approach that included self-report, peer report, and wearable sensor technology (Radio Frequency Identification Devices, or RFID). RFID has been shown to reveal social dynamics among children during school recess in an objective and unobtrusive manner [1, 35, 161].

With this set-up, first, we aimed to determine the levels of physical and emotional connectedness (Figure 4.1). RFID data revealed children's physical connectedness

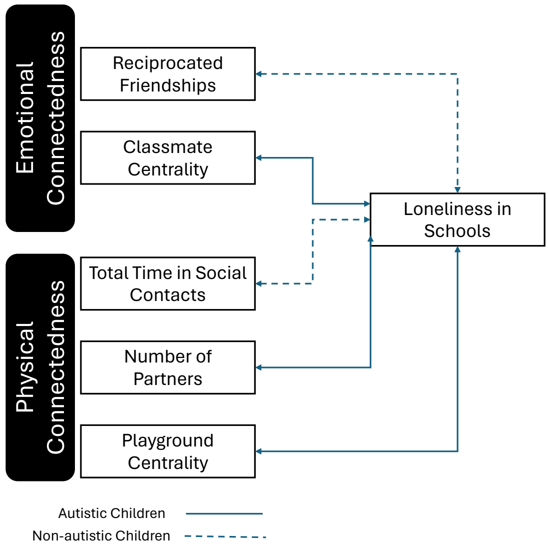


Figure 4.1: Overview of the study variables and hypotheses. Solid lines represent the hypotheses for the autistic group; dotted lines represent the hypotheses for the allistic group. The double-headed arrows denote the negative correlations expected between actual and felt social connectedness.

on the playground during school recess. From the RFID data, we computed (i) amount of time spent in face-to-face social contacts, (ii) each child's number of contact partners, and (iii) their level of connectedness to the entire playground social network (i.e., the degree of "centrality" that reflects how physically close each child was to all the other peers in the playground social network). For emotional connectedness, peer nominations were used to measure reciprocity in friendships, and each child's relative level of connectedness to the class social network (i.e., the degree of centrality that reflects how all the other classmates in the class network were emotionally available to each child). By including the centrality measure, we took into account pupils' connectedness to the larger peer network. Given that autistic pupils may value their social connections differently from their allistic peers [29, 142], we considered both the stronger and weaker social connections. We expected autistic pupils to be less connected than allistic pupils, as measured by peer nomination and by objective data collected in RFID sensors at recess, although this prediction was based on prior studies including autistic children in mainstream school settings. We also expected to find significant variance among individual scores, on all measures.

Second, we aimed to understand how children felt about their social connectedness at school. Therefore, we examined feelings of loneliness in school to understand the extent to which children were unsatisfied with their current level of social connectedness; and investigated the extent to which children's loneliness was related to their physical and emotional connectedness with peers, according to the above measures. We expected no difference in the levels of loneliness between autistic and allistic pupils, as the pupils in this study were still in primary school years [136,147,148]. Furthermore, previous studies with relevant age groups showed that more loneliness was associated with less intimate and less positive peer relationships among allistic pupils, whereas, for autistic children, more loneliness was associated with a lack of opportunities to safely be part of the school activities [29,142,157,159]. Therefore, we expected that allistic pupils would feel less lonely when they have more reciprocated friendships, and/or when they spent more time in face-to-face contacts; and autistic pupils would feel less lonely when they were included in school social activities more often, e.g., when they had contact with more peers during recess and/or occupied a more central position within their peer groups. However, this part of the hypothesis was exploratory in nature.

4.3 Methods

4.3.1 Participants

All pupils in this study attended two primary schools for special education in the Netherlands (School A and School B). Autistic pupils (n = 47) and their allistic classmates (n = 52), aged 8 to 13 years, were recruited to participate (M_{age}= 10.84 years, SD = 1.21; 34 girls and 65 boys). Nineteen (40%) of the autistic pupils had additional diagnoses related to psychiatric/behavioral conditions, such as attention deficit hyperactivity disorder (ADHD). Among the allistic pupils, 18 (35%) had a diagnosis of psychiatric/behavioral conditions that did not involve autism (see Table 4.1 for more details about the distribution of these diagnoses). The autistic group was younger ($t(96) = 5.41, p < .001$) and had fewer girls and more boys ($\chi^2 = 22.30, p < .001$) than the allistic group. Note that specific data on socioeconomic status were not recorded.

In the Netherlands, special education is divided into four clusters (1: low vision; 2: serious communication difficulties, e.g., hearing loss or language disorder; 3:

cognitive/physical disabilities or a chronic illness; 4: psychiatric or serious behavioral difficulties, e.g., autism, ADHD, and/or oppositional defiant disorder (ODD)). In this study, both autistic and allistic children were recruited from two Cluster 4 schools. Both schools also accepted several pupils without a specific diagnosis who switched from mainstream schools due to difficulties in adjusting to the pace of learning, class size, and peer interactions, and their need to receive extra care and support.

The special education setting provides activities that are more structured and predictable and gives more personal attention and specialist support to individual students. Both schools are members of the same private educational organization, using similar teaching methods and structuring their school activities and rules similarly. During recess, children shared the playground with peers from their grade; teachers supervised the recess time but did not intervene in the activities unless necessary.

Before a child can be admitted to a Cluster 4 school, the receiving school must request a declaration of admissibility from the regional education council (i.e., the governmental organization responsible for the management of education in the region). The council is obliged to be advised by at least two experts (from a committee of remedial educationalists, child psychologists/psychiatrists, social workers, and doctors) to verify the condition and issue an admissibility statement. Based on this system, in this study, we asked for the diagnosis information from parents and confirmed these with their teachers according to the school documents.

This study was part of a larger-scale research project that examines different aspects of social participation and inclusion of autistic children in schoolyards [1, 2]. Guardians of child participants signed informed consent forms prior to test procedures. The study protocol and informed consent form were approved by the Psychology Research Ethics Committee of Leiden University.

4.3.2 Measures and Procedures

Children completed self- and peer report questionnaires on a tablet, accompanied by either their teacher in the classroom or an experimenter in a separate room in school. Before filling out the questionnaires, they were presented an instruction video on the tablet, which described the purpose of the study, instructed how to fill out the questionnaires, and showed that they can ask questions. Teachers and experimenters were instructed to only provide support when necessary, i.e., when children asked questions or when they appeared to have misunderstood the questions.

Sensor data were collected from each child on four occasions on two school days, during morning and lunch recess time on both days, each lasting 11 to 30

Table 4.1: Background characteristics of the participants.

	Autistic	**Allistic**
N	47	52
Age, years, mean (SD)	10.20 (1.00)	11.38 (1.14)
Gender, n (%)		
Girls	5 (11)	29 (56)
Boys	42 (89)	23 (44)
School distribution, n (%)		
School A	6 (13)	46 (88)
School B	41 (87)	6 (12)
Playgroup allocation, n (%)		
Lower grades (5-6)	27 (57)	27 (52)
Higher grades (7-8)	20 (43)	26 (48)
Additional psychiatric/behavioral conditions, n (%)		
None or unknown	6 (13)	35 (67)
Autism only	24 (51)	-
Attention Deficit Hyperactivity Disorder (ADHD)	17 (36)	16 (31)
Developmental Language Disorder	2 (4)	1 (2)
Oppositional Defiant Disorder	0	1 (2)

minutes (M = 18.97 minutes; SD = 6.62). Before recess, all pupils were given a belt they wore on their waist, on which a RFID tag was mounted, facing front. Pupils were explained that they could take off the belt when they were not comfortable with it, but only 1-2 children in 2% of the break sessions did that. They wore the belt throughout recess on the playground, and returned the belt when the recess ended. Teachers on the playground were also given a sensor belt, although in this study only the social contacts between pupils were considered. During recess on the playground, children were not given specific instructions regarding where or with whom to play.

4.3.2.1 Measures of Physical Connectedness via Wearable Proximity Sensors

OpenBeacon RFID tags are proximity sensors by means of Bluetooth, registering face-to-face contact between pupils on the playground during recess. RFID is an unobtrusive and objective measure that allows for quantifying spatial proximity between children in their daily school settings, continuously throughout a recess. It does not intervene children's behavior and ensures ecological validity [76]. Previous research has proven the accuracy and specificity of the RFID technology, showing that the social contacts detected by the RFID tags corresponded to video observations and self-reported amount of social contacts in both adults and children [35,77]. Moreover, RFID tags have been previously used in children with autism and hearing loss [7].

Pupils on the playground were each given a RFID tag mounted to a belt, facing front. When two children, whilst wearing an RFID tag, were facing each other and within a distance of 1.5 meters, the tags detected Bluetooth signals and passed on the data to the signal-receiving base station, which then registered a social contact. RFID tags could detect multiple contacts simultaneously. To compensate for unintended interruptions, contacts with interruptions shorter than 35 seconds were registered as one single contact [1,161]. The receiving station captured RFID signals from an area covering 15 m^2 four times per second [76]. It was located on the school playground at a pre-determined location to ensure maximized detection range. To ensure a fair comparison between all children, we only considered the RFID records of a participant when the participant was detected by the receiving station for at least 50% of the recess time. Otherwise, the data for that participant in that specific recess session were excluded from further analyses. Sensor data points from the four measurements (two breaks x two days) were averaged for further analyses. We included three variables derived from social networks detected by the RFID (Figure 4.2):

- *Total time in social contacts* indicated the proportion of time a participant spent in face-to-face contacts during recess. This was calculated by dividing the total duration of time spent throughout all contacts by the total duration of time that a participant was detected by the receiving station, in a specific recess session.

- *Number of contact partners* indicated the number of different peers that a participant had social contact with during recess. This was calculated by dividing the total number of contact partners by N-1, where N denotes the total number of pupils detected on the playground in a specific recess session.

- Lastly, we calculated *playground closeness centrality* from the RFID data to examine each pupil's connectivity to all the other peers on the playground. The closeness centrality was computed according to the following formula [162]:

$$C(u) = \frac{(n-1)}{\sum_{v-1}^{n-1} d(v,u)} \tag{4.1}$$

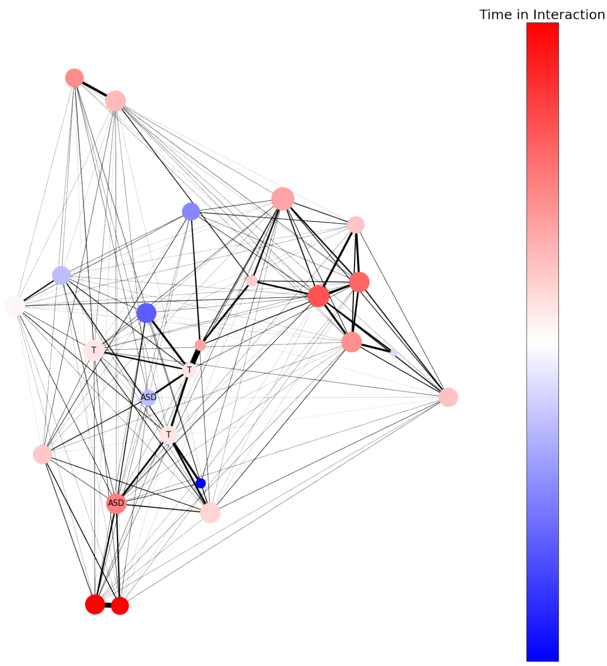


Figure 4.2: Visualization of a sample social network detected by the proximity sensors (RFID) during one school recess session. Each node represents an individual on the playground. Autistic children are labeled with "ASD"; teachers with "T". The color of the nodes denotes total time in social contacts during this recess session; warmer colors (red versus blue) suggest longer time. The thickness of the edges between two nodes denotes the duration of dyadic contacts; thicker edges suggest two notes/children having longer contacts with each other. The nodes are positioned in accordance with their centrality.

where $d(v, u)$ denotes the shortest-path distance between Child v and Child u, and n denotes the total number of participating children on the playground. Here, we further weighed this score by dyadic contact time measured by the

RFID, to reflect how closely a participant was connected to the network, both in terms of the shortest-path distance to reach all the other peers in the network, and the time spent with them.

Dyadic contact time refers to the duration of contact between two children, normalized by the total duration when both children were detected by the receiving station, in a specific recess session. The inverse of this value was used as the weight; thus, longer dyadic contact time led to a smaller weight (i.e., a shorter path), which led to a higher playground closeness centrality. Therefore, this playground closeness centrality measure combined both the time and the number of partners in each social contact, reflecting individual children's relative position in their playground social network [162, 163]: a higher playground closeness centrality in this study thus indicated that the participant was more physically close to all the other peers on the playground than those with a lower closeness centrality.

4.3.2.2 Measures of Emotional Connectedness via Peer Nomination

To examine reciprocated friendship, peer nominations were obtained by asking each child to write down the names of their best friends in school. They could provide maximally five names. This limitation of maximally five names ensured that we obtained the stronger social connections that can be recalled by the children, following [164]. From these nominations, we derived the number of nominations given by each participant (i.e., outdegree in social network analysis) and the number of nominations that were reciprocated (i.e., bi-degree; the participant nominated a peer as a friend, and the peer also nominated back). We computed the degree of reciprocated friendships by dividing bi-degree by outdegree.

To examine each pupil's connectedness to their larger peer network in the class, each participant was presented with a list of classmates, and they answered the extent to which they liked to play with each classmate (i.e., "yes", "sometimes", "no", or "I don't know"). Based on peers' ratings, we calculated a *classmate closeness centrality score* for each participant. A higher classmate closeness centrality score indicated that the participant could approach all the other peers in the class more easily – more likely to be 'liked' – than those who with a lower centrality [163]. When Child v receives a "yes" or "sometimes" rating directly from Child u, the two children are seen as having a connection with an one-unit distance. Based on all the classmate ratings in each class, their centrality scores were computed using the same formula described above. Here, we treated both "sometimes" and "yes" answers as indicating being part of the social network, but "sometimes" answers were weighted with a distance of 1, while "yes" answers were weighted with a distance of 0.5 (i.e., a shorter path). Only the ratings given by peers were considered.

4.3.2.3 Loneliness in School via Self-Reports

Children's Loneliness Scale (CLS) was used to assess levels of loneliness in school in terms of children's dissatisfaction with social connections [165] (validated in Dutch-speaking children: [166]). This is a self-report consisting of 24 items, rated on a 5-point scale (1 = not at all; 5 = always). Six items that were positively formulated were reverse scored, thus higher scores indicating higher levels of loneliness. Eight items were control items about children's hobbies and preferred activities, and were excluded from further analyses. Internal consistency was good (Cronbach's $\alpha = .87$ for all children; $.87/.88$ for autistic/allistic pupils).

4.3.3 Statistical Analyses

Closeness centrality from peer reports was computed using igraph within R [167, 168]. Variables derived from the RFID data were preprocessed and computed using Python 3.9 [169]. The NetworkX 2.6.3 Python package was used for visualization. Statistical analyses were performed using SPSS version 27.0 (SPSS Inc., Chicago, IL, USA).

First, to examine the extent to which autistic and allistic children differed in their connectedness in social networks (physical and emotional connectedness) and loneliness, a series of Mann-Whitney U tests were conducted. Rank-based nonparametric tests were used due to the presence of outliers in the variables of loneliness, playground closeness centrality, and classmate closeness centrality. To assess whether the variance among individual scores was equivalent between the two groups, Levene's tests of homogeneity of variances, based on the deviation from the median values, were used.

Next, to examine the extent to which feelings of loneliness were related to connectedness in social networks, partial Spearman's correlation tests were administered, controlling for age. Fisher's r-to-z transformation was used to examine the moderating role of autism diagnosis, comparing the strength of correlation between autistic and allistic children. To correct for multiple testing, the Bonferroni procedure was applied, and the significance level of the main analyses was adjusted to $p < \alpha/5 = .01$. Little's MCAR test showed that data were missing completely at random ($\chi^2 = 177.82$, $p = .111$). Thus, we handled missing data using the multiple imputation (MI) technique [170–172]. Ten imputations were performed. Given the age and gender differences between autistic and allistic pupils, and to reduce the effect of the school of origin, the inverse probability of treatment weighting (IPTW) procedure was applied [173–175]. IPTW is a widely used weighting method for adjusting confounding variables. Following the procedures proposed by Chesnaye et al. [174], first the probability of a participant being in the autistic vs. allistic group was computed into a propensity score, taking into account individuals' characteristics (i.e., age, gender, and school).

Next, a weight was assigned to each participant, computed as the inverse of the propensity score, through which potential confounds were balanced across groups. To avoid extreme weights that may bias the outcomes, the weights were stabilized by accounting for the proportion of autistic vs. allistic children, and extreme values beyond the 1st and 99th percentiles were truncated. Pooled and weighted results were reported. Correlations between all study variables are presented in Appendix A. Results based on raw data are reported in Appendix B. To understand the effect of participant heterogeneity, we also ran the analyses while excluding the autistic children with comorbidity and allistic children with a diagnosis (Appendix C), and compared social connectedness between autistic children with and without comorbidity (Appendix D).

4.3.4 Community Involvement

The overall objectives of the larger project were formulated in meetings with autistic self-advocates and researchers, as well as practitioners working with autistic individuals. Also, the new methodology involving sensing technologies was discussed with associations for promoting the interests of autistic people, school organizations, and governmental organizations, although they were not directly engaged in formulating the research questions addressed in this study.

4.4 Results

4.4.1 Levels of Physical and Emotional Connectedness and Loneliness

Table 4.2 shows the mean levels and standard deviations for the study variables. Regarding the observed levels of social connectedness as measured by peer reports and RFID, autistic children had fewer reciprocated friendships ($U = 1528.0, p = .002$) than allistic children. There were no group differences in total time in social contacts, in the number of contact partners, or in classmate/playground closeness centrality ($U_s \geq 1878.0, p_s \geq 0.145$). Regarding the levels of loneliness, no group

Table 4.2: Mean levels and standard deviations (SD) of social connectedness and the Spearman's correlations with loneliness (controlling for age)

		Mean (SD)			Correlation with loneliness		
	Range	Autistic	Allistic	U	All	Autistic	Allistic
Loneliness (total score[a])	16-68	33.38 (8.03)	37.03 (11.11)	1904.5	-	-	-
Physical connectedness							
Time in social contact[b]	0.03-1	0.62 (0.22)	0.63 (0.20)	1878.0	-0.06	0.30	-0.39***
Number of partners[c]	0.09-0.95	0.56 (0.15)	0.56 (0.15)	2002.5	-0.08	-	-
Playground Closeness centrality	0.02-0.11	0.08 (0.02)	0.08 (0.02)	1892.0	0.12	-	-
Emotional connectedness							
Reciprocated friendships[d]	0-1	0.39 (0.25)	0.48 (0.24)	1528.0**	0.04	-	-
Classmate Closeness centrality	0.47-2	1.05 (0.34)	1.12 (0.34)	2091.5	-0.10	-0.36**	0.08

Correlation coefficients for separate groups are reported only when Fisher's r-to-z transformation showed a significant difference in the strength of correlations between the group; otherwise, the correlation coefficients for the entire sample are reported.
[a] Highest possible total score is 80.
[b] Corrected by the total time when the child was detected.
[c] Corrected by n − 1, where n is the total number of children on the playground.
[d] Calculated as a degree by dividing the number of reciprocated nominations by the number of outgoing nominations.
$^{**}p \leq .01$.
$^{***}p \leq .001$.

differences were noted in the levels of loneliness ($U = 1904.50, p = .181$). Tests of homogeneity of variances showed that variances for the loneliness scores were not equal between the two groups ($SD_{autistic} = 8.03 < SD_{allistic} = 11.11, F(1, 93) = 5.04, p = .027$). For the other variables, the variances were equivalent across the groups, Fs < 1.60, ps > .208.

4.4.2 Relations between Social Connectedness and Loneliness

Higher classmate closeness centrality was related to lower loneliness only in autistic children ($rho = -.36, p = .004$), not in allistic children. More time spent in social contacts was related to lower loneliness only in allistic children ($rho = -.39, p < .001$), not in autistic children. No other significant correlations or group differences in correlational strength were noted (Table 4.2).

4.5 Discussion

The present study aimed to examine how autistic children construed their social connectedness in school, compared to their allistic classmates. In line with our expectations, autistic pupils in this study had fewer reciprocated friendships than allistic pupils. Unexpectedly, autistic and allistic children were similarly connected to their peers, in terms of time spent in social contacts and number of interaction partners during recess, and their centrality in classmate and playground networks. The levels of loneliness experienced by the two groups at school did not differ. However, the factors related to their loneliness did differ: while allistic children felt lonelier when they spent less time in physical social contact during recess, autistic children reported higher loneliness when they were less central, i.e., less liked as a classmate to play with, in the classmate network. No other relations were noted.

Unlike many previous studies that reported autistic children were less connected to their peers than allistic children [141, 143, 144, 148–150], in this study we found comparable outcomes in most aspects of physical and emotional connectedness, including total time in social contacts, number of contact partners during recess, and their centrality in peer networks. That is, we observed that autistic children were in social contacts at the group level to the same extent as their allistic peers.

It could be argued that this positive picture was likely a result of the special education setting where we collected data. That is, compared to settings in mainstream education, autistic pupils in special education are usually not the only ones with a diagnosis; class sizes in special education are smaller; school activities including recess are more structured; and teachers are better equipped with skills to identify problems and support and facilitate positive social dynamics among children [176–179]. In fact, in our sample, autistic children were the majority in several of the classes. In this context, autistic children have more opportunities to meet their autistic peers and other peers with special needs. When it is recognized that all pupils have their unique needs, being "different" with a diagnosis and social difficulties may be less of an issue for joining peer activities [180]. A more structured recess in special education settings may also allow more face-to-face contact to be facilitated and thus be detected. Though beyond the scope of the current study, our findings suggest that the school climate and how the school environment is organized may play an important role in children's social participation beyond individual children's diagnosis and social skills. Several prior studies have also shown that autistic children become more socially engaged when the school playground is adapted to provide more equitable opportunities that also address autistic children's needs and capacities, e.g., by reducing noises and improving acoustic to lower overstimulation commonly encountered by autistic pupils in school settings; by offering more structure (e.g., by making different compartments with different functions); and by addressing different sensory needs (e.g., to set up different sensory zones and transition between the zones [181–184]).

Our use of sensor technology may also have contributed to such findings. The present study showed the first attempt to capture social dynamics between autistic children and their peers throughout recess, in a naturalistic setting, objectively and unobtrusively (see [1,76] for more information about this methodological approach). This method returns objective and richer information regarding children's group dynamics, complementing methods in previous studies (e.g., systematic observation), which can be constrained by observation timeframe, observer bias, and the coding scheme.

However, it seems that many autistic children were not considered friends by peer group members and had fewer reciprocated friendships than allistic peers. The peer-report measure for reciprocated friendships was the only social connectedness indicator in this study that required active responses from peers, and it denoted the reciprocity perceived by peers, towards a specific child. To receive more reciprocated friendship nominations in a free-recall task, a child has to be prominent enough in a network for the other peers to select them as friend. In such a scenario, it appears that autistic children were more often overlooked by their social group members.

The question then is, how did the autistic children feel, in light of their lack of connectedness? Our results showed that the degree of reciprocated friendships was unrelated to feelings of loneliness, i.e., there was no mismatch between the reciprocated friendships that autistic children desired and perceived. Possible explanations for this include that (i) autistic children were not particularly aware of reciprocity in their social connections; (ii) they did not care about or want many connections, or acted like they did not care for self-protection; or (iii) they did belong to a social group, despite not having many apparent connections [148]. Adding to the study by Chamberlain and colleagues [148], our results seem to substantiate the third explanation. In this study, the large majority of autistic children did have at least one peer who liked to play with them ($n = 44$; 94%), and they also experienced that they had at least one friend ($n = 38$; 81%), which might meet their needs to know that they are not alone. Moreover, in autistic children, more loneliness was related to being less liked in a classmate network (i.e., lower classmate closeness centrality, an aspect of emotional connectedness), whereas in allistic children more loneliness was related to spending less time in social contacts during recess (an aspect of physical connectedness). Apparently, for these autistic children, being liked as part of a social group and experiencing group-level emotional connectedness is closely related to their feelings of loneliness. It is thus likely that in autistic children, social connectedness is not evaluated based on dyadic contacts, but beyond that, based on the extent that one feels accepted by a group. When they do feel they belong to a social group, the risk of them feeling lonely might decrease.

These results highlight the importance of looking into individual differences and widening the possible definition of social relationships. Different individuals could find different features of social connection to be valuable. While spending time together may be seen as a relationship goal by many people, it may not always be the case for others. Actually, in autistic children, total time in social contact during recess had a positive, rather than negative, correlation with loneliness (rho = .30, although not reaching the significance level after the Bonferroni correction). Possibly, having to stay in face-to-face interaction with others could cause stress, anxiety, and exhaustion in autistic children, as they must constantly attend to social cues that they may not fully understand, and/or some may even feel the need to camouflage or hide their social difficulties so as to "fit in" [185–187]. These challenges can be further aggravated in adolescence due to the even less structured environment – moving from one classroom to another, going to canteens, and having no supervision during recess – which may underly autistic adolescents' elevated levels of loneliness in school as reported by previous studies [29, 136, 159]. Nevertheless, despite the possible differences in how social features were viewed, the fact that autistic children experienced loneliness in school at varying levels, to an extent that was comparable to that of allistic children, shows that autistic children do value social connection and are aware of unfulfilled desires for social relationships, like their allistic peers.

4.5.1 Limitations and Future Research

This study was among the first to examine physical and emotional connectedness in autistic children and their relations with loneliness in school, using a multi-method approach that accounted for social dynamics over an entire recess session. Yet, some limitations should be taken into account, and some caution is due when interpreting the results.

First, as mentioned earlier, all participants were from special education schools and many allistic pupils also had a diagnosis, although not autism. We were thus able to investigate what the social situation is like when autistic pupils are substantially represented in a class network. This may explain why the measured levels of social connectedness were largely similar across the groups, while previous studies in mainstream schools (where usually only one or two autistic pupils were present) showed lower levels in all aspects [141, 144, 148–150]. Future studies could explore the extent to which these outcomes are generalizable to mainstream settings, and how the findings from special education settings may inform inclusive practices in other settings. Future research is also required to understand how the school climate may vary with different school policies for inclusion and diversity, and how an improved school climate may further enhance children's social wellbeing. Focus group interviews could be organized with children belonging to different groups, to take a variety of views into account.

Here, our finding regarding the importance of group-level emotional connectedness – a sense of belonging to a group – among autistic pupils requires special attention. This may become increasingly difficult to achieve as their social challenges intensify during adolescence. This may (partially) explain the higher levels of loneliness found in autistic adolescents when compared to their allistic peers [29, 136, 142, 157–159]. Moreover, this may imply that school policies and practices could be especially influential for the school life of these pupils.

Second, half of the autistic children in this study had comorbidity (besides their diagnosis of autism; see [188] for a review on comorbidity in autism), and 35% of the allistic children had at least one diagnosis. Given the Cluster 4 special education setting, the majority of the participants were not neurotypical. These sample characteristics should be taken into account, as they influenced the social dynamics. Future studies are encouraged to confirm whether the patterns we found for the allistic group is specific to this sample that included many pupils with ADHD and several pupils who transitioned from mainstream education. However, we included all children, regardless of their diagnosis, in the study because our main goal was to examine the effect of autism on social connectedness. Our findings highlighted how pupils with and without autism may have different social needs, which should be considered when providing support.

Notably, in our follow-up analyses, where we excluded autistic children with comorbidity and allistic children with a diagnosis, we confirmed the same differential patterns (Appendix C). However, after those children with comorbidities were removed, autistic and allistic children no longer differed in the number of reciprocated friendships. Further inspection added that autistic pupils with comorbidities had fewer reciprocal friends, while contacting more partners – hence spending shorter time with each partner – during recess, compared to autistic pupils without any comorbidities (Appendix D). While these outcomes might be affected by the smaller sample size, it is likely that comorbidity could put autistic children in a more vulnerable position, regarding the formation of close relationships. Future studies are needed to further understand the needs and wishes for social connectedness of children with comorbidities or other diagnoses, for creating a school environment where all children are respected.

Third, the two groups differed in several aspects, besides their diagnoses. The autistic group was older and featured fewer girls than the allistic group. While the

IPTW method can balance the characteristics in the samples with reliable outcomes, effectively reducing the impact of confounding effects, there could still be potential biases in the results that we overlooked. Despite the fact that the two schools we selected were managed by the same educational organization and organized school activities in a largely similar way, there might be factors in the two schools that influenced our results. To further assist interpretation of our findings, results based on raw data were also presented (Appendix B). Whether weighted or not, our results consistently revealed group differences in reciprocated friendships, and in the relation between social contact time and loneliness in allistic children, but not in autistic children.

Fourth, some limitations in our data collection should be noted. In this study, the Children's Loneliness Scale [165] was used to examine levels of loneliness in terms of children's unmet needs of social connections. Yet it should be noted that this measure has not been formally validated among autistic children. Prior studies have shown that autistic children may define loneliness differently from their allistic peers: they tended to focus on the dissatisfaction of social connections, whereas allistic children more often also mentioned the associated negative affect [29]. This possible discrepancy should thus be considered in future studies, and we also call for further research to validate existing loneliness measures in autistic youth separately. Also, peer nominations were limited to the school setting, yet it is likely that children also have connections outside of school, e.g., in the neighborhood. Further, the proximity sensors, i.e., the RFID badges, captured only face-to-face contacts within 1.5 meters. Social contacts are not always on a face-to-face basis, or within such a close distance. Children in a playground may play together side-by-side (which is often observed in autistic children [189, 190] or talk to each other from a distance, but those interactions could be largely missed with the current configuration. Moreover, these detected social contacts may not necessarily reflect social engagement, as pupils may be in close proximity but not involved in a joint activity. The governmental measures such as social distancing in response to the COVID-19 pandemic might also affect the social dynamics between children to some extent, although at the time of our data collection, no constraints were imposed.

4.6 Conclusions

Most children go to school on a daily basis. Understanding how they feel and how to promote their wellbeing in school is of utmost importance. Such knowledge, however, is limited in the literature, especially regarding children with special needs, who compose at least 10% of the student population [191]. Our findings provide evidence that loneliness in school may be construed differently by autistic and allistic children, although the levels of loneliness were comparable in the two groups. For these autistic children, feelings of loneliness may go beyond face-to-face interactions. Rather, being liked as part of a peer group was key.

Our findings call for further investigations that examine individual differences in social connectedness; and for school-based interventions that move the focus from individual children's social skills, for the purpose of "fitting into" peer activities, to adapting the school environment so it promotes inclusion. Understanding relevant differences in children's needs could well lead to more effective design for a welcoming school climate. This in turn could increase social wellbeing of not only autistic children, but all children at school.

CHAPTER 5

SiamCircle: Trajectory Representation Learning in Free Settings

The contents of this chapter are based on the following submitted study:

- M. Nasri, M. Baratchi, A. Koutamanis, C. Rieffe, “SiamCircle: Trajectory Representation Learning in Free Settings”, IDA, 2025.

Abstract

Trajectory representation learning (TRL) is an intermediate step in handling trajectory data to realize various downstream machine-learning tasks. While most previous TRL research focuses on modeling structured movements in large-scale urban spaces (e.g., cars or pedestrians on streets), this paper focuses on a more challenging scenario of modeling free movement in small-scale social spaces (e.g., children playing in a schoolyard). We present SiamCircle, a novel contrastive learning approach for TRL uniquely designed to process raw trajectories without additional feature extraction to prevent information loss. SiamCircle adopts a Siamese network with Circle Loss to learn trajectory embeddings. Furthermore, SiamCircle employs a data augmentation process to enable self-supervised learning and enrich the input data to address the limited access to high-quality data and ground truth. We evaluate the performance of SiamCircle in downstream tasks using three classes of metrics, namely trajectory ranking, trajectory similarity, and clustering performance, using collectively nine evaluation metrics. Using an ablation study, we explored the impact of different neural network components and loss functions on the model's performance. Accordingly, we selected a 2-D convolutional design with Circle Loss as the best-performing model. In a comparative study, we compared our model against three other baselines. We observed up to 19% improvements in trajectory ranking tasks and achieved the highest average rank in trajectory similarity and supervised clustering tasks.

5.1 Introduction

In the wake of rapid growth in wearable technologies, people generate large amounts of movement data using location-aware devices like sports bands and smartwatches. Various organizations address existing challenges by collecting and analyzing this data, known as spatio-temporal movement data or trajectories [192].

In a broader context, this movement data is captured in mainly two settings: *structured settings* and *free settings*. In structured settings, movement is recorded in areas influenced by spatial features or regulations, e.g., road networks or driving policies. This setting often exhibits periodic patterns, such as commuting to work on weekdays. Data from structured settings are valuable for applications like traffic forecasting. Conversely, in free settings, individuals move without restrictions, resulting in unstructured trajectories, which often occur in constrained environments, such as students' movements in a schoolyard or athletes' movements in sports. This data is useful for analyzing micro-level behaviors, including social interactions and

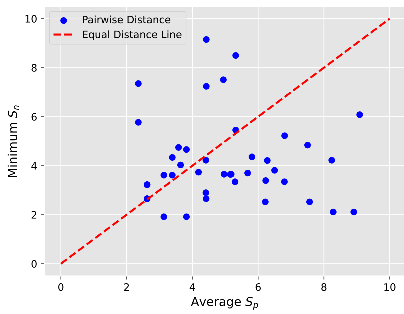


Figure 5.1: Average pairwise distances between a trajectory and trajectories from the same group (S_p) versus from different groups (S_n). Many trajectories have smaller S_n than their S_p, i.e., some samples are positioned below the red line.

group dynamics.

In the context of trajectory analysis in structured settings, trajectory representation learning (TRL) has attracted extensive research attention in various machine learning tasks such as trajectory flow forecasting [192] or pair-wise similarity computation [193]. Yet, one largely unsolved challenge is effectively designing a TRL framework for movements occurring in free settings. Developing TRL models in free settings is more challenging than in structured settings due to (i) irregular patterns and movements formed in the absence of typical urban features, (ii) limitations in data acquisition systems leading to imperfect location data and complicating trajectory analysis, and (iii) privacy concerns restricting the collection and sharing of high-quality data.

To understand these kinds of challenges, we present the average pair-wise distribution of trajectories in the ETH dataset [37], which captures trajectories in a free setting, annotated by group membership. Figure 5.1 presents the pairwise Euclidean distances between trajectories of the same group, i.e., positive samples (S_p), and different groups, i.e., negative samples (S_n). The red line shows where these two distances, i.e., S_p and S_n, are equal. In general, one might assume that, given individuals' tendency to walk closer to the group they belong to, all samples within the same group should have smaller S_p values compared to their S_n (i.e., all samples to be positioned above the red line). However, here we observe that a large number of samples maintain higher average S_p values than their minimum S_n values being positioned below the red line. Many representation models, particularly auto-encoders [194], which learn embeddings without explicit differentiation between similar and dissimilar samples, may struggle to resolve this complexity. This will, in turn, limit their performance in downstream tasks.

Furthermore, recent studies in modeling movements in structured settings implement spatio-temporal pre-processing methods such as behavioral feature extraction [194], grid-cell projection [195], and graph embeddings [193]. While the results are promising, their application to trajectories collected in free settings may be limited due to the potential information loss during the pre-processing. Besides, their performance highly depends on the choice of hyperparameters, such as cell size. In a free setting, finding the optimal grid size is more challenging, especially in the presence of imprecision due to noise.

This paper presents a robust trajectory representation learning framework called SiamCircle, which is directly applicable to raw trajectory data collected in free movement settings. SiamCircle consists of three main components: (1) a data augmentation process to generate the augmented trajectories accounting for imperfections in trajectories and enriching data sets with limited sample size, (2) a neural network applicable to raw trajectory data that maximally exploits spatial and temporal similarities among similar trajectories while amplifying differences between non-similar ones, eliminating common information loss during feature extraction, (3) a loss function uniquely designed with our learning framework to learn representative features through the training process. To our knowledge, this is the first study proposing a TRL framework using raw trajectories collected in free settings. Overall, this paper makes the following contributions:

- We develop a framework to create augmented trajectories that simulate similar movements and imperfections in data. Moreover, this framework generates a ground-truth dataset that annotates the trajectories based on their similarity class and enriches datasets with a limited sample size.
- We propose a unique deep neural network model with triplet-based Circle Loss to learn trajectory representations directly from raw trajectories to limit information loss commonly occurring during the feature extraction.
- We demonstrate the performance of the proposed framework by conducting an ablation study and a comparative study on five trajectory datasets collected in a free setting using nine evaluation metrics in downstream task using three classes of metrics, namely trajectory ranking, trajectory similarity, and clustering. We compare the performance of SiamCircle to three other baseline models.

The remaining part of the paper is organized as follows. In Section 5.2, we discuss related work. Section 5.3 presents the problem statement, and Section 5.4 explains the details of our proposed method. The experimental setup and results are presented in Section 5.5 and Section 5.6, respectively. Finally, we summarize the paper and discuss future research directions in Section 5.7.

5.2 Related Work

Several studies proposed TRL models primarily in structured settings like urban traffic forecasting and detecting transportation modes [196, 197]. Recent studies incorporated external data, such as road networks and urban structures, into TRL models [198, 199]. While these models show promise, they become impractical for data collected in free settings, especially in the absence of external information. Moreover, several studies adopt feature engineering techniques to transfer raw trajectories into an abstract representation before training their model [193, 194]. Trajectory2vec [194] proposed a manual feature extraction algorithm to extract space- and time-invariant characteristics of trajectories, e.g., change of rate of turns. This model adopted a sequence-to-sequence auto-encoder to generate a deep representation of trajectories from extracted features.

Moreover, TRL via Contrastive learning has been employed in various machine learning problems in natural language and image processing and, more recently, in the context of spatial problems [193, 200, 201]. Contrastive learning is especially beneficial when using noisy and imperfect datasets with limited sample sizes. Fan Z. et al. [201] proposes a contrastive learning approach powered by Triplet loss [202] in combination with a sliced encoder network for a user re-identification problem based on trajectory data. Their method, S-BiLSTM, learns the robust part of individuals' trajectories (segmented in 24 hours). Such a 24-hour sliced design is not applicable for movements in free settings where trajectories do not exhibit regular patterns. TrajCL [193] is another contrastive learning model including a dual-feature multi-head self-attention-based encoder with InfoNCE loss [203] using trajectories in urban areas. TrajCL proposed a pointwise trajectory feature enrichment method to extract structural and point features to train their model. The optimization process used in both loss functions, Triplet loss and InfoNCE loss, is rigid, as their loss calculations give pairs with various similarity ranges an equal pre-defined margin.

The present study uses a Siamese network architecture with Circle Loss to learn trajectories' representative features designed explicitly for free movements in constrained environments (i.e., free settings). Our model uses raw trajectories without any feature extraction techniques as opposed to earlier work [193, 194] to limit the unnecessary loss of data and oversimplification in free settings. We further implement a dynamic penalty scheme based on the degree of similarity to obtain a

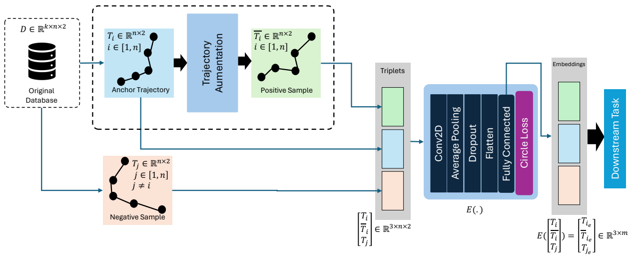


Figure 5.2: An overview of SiamCircle: Data Augmentation Process to create the [Anchor, Positive, and Negative] triplets, Siamese Network with Circle Loss to generate the embeddings of the given raw triplets.

sustainable optimization process and high-quality embeddings to address the rigidness problem mentioned above in earlier works (see, e.g., [193, 194]).

5.3 Problem Definition

Our trajectory modeling problem is based on a given trajectory dataset D, where each trajectory $\{T_i \in D | T_i = \{(x_1, y_1), \cdots, (x_n, y_n)\}\}$ is a sequence of two-dimensional points recording the trace of the moving object i with a fixed size n where (x_n, y_n) is the n^{th} location of the object.
We are interested in learning an embedding function $E(\cdot)$ that represents trajectory T_i as $E(T_i) = T_{i_e} \in \mathbb{R}^m$ ($m \ll n$ is the dimension of embedding space) such that $f(T_{j_e}, T_{i_e})$ is minimized for any pair of the object i and j moving along the same underlying route, and, conversely, maximized for any pair of objects i and j moving along different underlying routes. Here, $f(\cdot, \cdot)$ is a distance function (e.g., Euclidean distance or other distance measures).

5.4 Methodology

This section discusses our proposed method for TRL. First, we explain the trajectory augmentation process. Next, the neural network design is presented. Lastly, we present the loss function used in the proposed framework. An overview of our method is presented in Figure 5.2.

5.4.1 Trajectory Augmentation

In this section, we propose a trajectory augmentation method that adopts meaningful transformations to generate augmented trajectories from original ones. Previously proposed trajectory augmentation methods involve applying pre-defined transformations to the original data, for instance, by truncating or point masking [193, 195]. These transformations often modify the trajectory length, making them only applicable in combination with manual feature extraction methods to create fixed-length sequential data for neural network models. However, as we aim to utilize raw trajectories with no feature extraction, we propose a unique trajectory augmentation approach where the trajectory length remains unchanged while meaningful transformations are applied. Our trajectory augmentation process serves three main purposes: (1) simulating similar movement trajectories by distorting original trajectories in different scales (i.e., additive noise) to account for different levels of similarities, (2) enriching limited-size datasets by adding augmented trajectories to the original dataset used purely for training. Thus enabling a self-supervised training process for the neural network model, and (3) simulating the practical imperfections found in location data acquisition systems in the form of large-scale noise (i.e., large spatio-temporal displacements or jumps), aiming to train a model that is robust to issues such as noise and other imperfections in the data. To achieve these goals, the augmented trajectory $\bar{T}_i$ is created by applying pre-defined transformations to each original trajectory T_i of a moving object i, as formulated in Equation 5.1:

$$\bar{T}_i(\bar{x}, \bar{y}) = (x_i + x_i', y_i + y_i'), \qquad \begin{cases} x_i' = \mathcal{U}(l_x, h_x).d_s, \\ y_i' = \mathcal{U}(l_y, h_y).d_s, \end{cases} \tag{5.1}$$

where $\bar{T}_i(\bar{x}, \bar{y})$ is the augmented trajectory of moving object i. x_i is the x coordinate of original trajectory T_i. x_i' is the distortion vector across x coordinates. $\mathcal{U}$ is the uniform distribution bounded between lower bound $l_x = -\sigma(\Delta x_i)$ and higher bound $h_x = \sigma(\Delta x_i)$, in which $\sigma(.)$ is the standard deviation. Δx_i is the positional differences across x coordinates. d_s is a distortion scale randomly selected among the given scaling range. In all the mentioned annotations, replacing x with y gives the same definition across the y coordinate (instead of x). In our experiments, two types of d_s are used, namely *additive noise* and *jumps*. While additive noise applies to all coordination of the original data, jumps only apply randomly to k number of samples. Another difference between additive noise and jumps is the scale of the distortion. In the additive noise, d_s includes a lower distortion scale to create similar movements, satisfying aims (1) and (2). In the jumps, d_s includes a larger distortion scale to create imperfections in trajectories, satisfying aims (2) and (3).

Thus, the semi-synthesized trajectory, $\bar{T}_i(\bar{x}, \bar{y})$, creates one variation of the object i in the database per distortion scale. In Section 5.4.2, we will discuss our proposed neural network model.

5.4.2 Siamese Network Architecture

This section discusses the design of our proposed neural network model. The unique aspect of this design is the model's ability to be directly applied to raw 2-D trajectories without implementing any preliminary pre-processing steps. The rationale behind this choice is to include as much information as possible without abstracting spatial details, e.g., trajectory griding, or extracting spatio-temporal features, e.g., speed. To achieve this goal, we adopted a Triplet-based Siamese network, which typically consists of three identical networks with shared weights [204]. The primary purpose of Siamese networks is to learn the similarity between inputs by comparing their representations. In this design, the model takes an anchor trajectory together with a positive sample (i.e., a trajectory from the same class as the anchor) and a negative sample (i.e., a trajectory from a different class than the anchor) as a triplet input, pass each sample through its identical branch, and hand over the generated embeddings to a triplet-based loss function. Specifically, our model contains two sections:

2D-CNN/Average Pooling/Dropout. Inspired by neural networks in time-series forecasting [205, 206], the 2-D convolutional layer is deployed as the first layer to extract feature maps by sliding the learning filters through the input. This section is further developed by an average-pooling layer, which reduces the noise and helps the network generalize better, and a Dropout layer to prevent overfitting. Overall, this section casts the trajectory into a more compact representation while retaining the most significant features in trajectories.

Fully-connected Layer. The outputs from the previous layer representing high-level embedding features are fed to a Fully-connected layer. This allows the model to learn the non-linear combinations of these features. The generated embeddings will be given to Circle Loss to allow the training process as detailed in Section 5.4.3.

5.4.3 Loss Function: Circle Loss

Most loss functions, including the triplet loss [202], directly incorporate the pairwise distances, i.e., s_n and s_p, into a similarity score. This optimization approach is rigid as it enforces equal penalty strength on each similarity score. Ideally, we would like to give greater emphasis (or penalty) when a similarity score significantly deviates from the optimum. To this end, Sun. et al. [207] proposed Circle Loss, which re-weights each similarity to highlight the less optimized similarity scores. The Circle Loss was initially proposed in computer vision with a unified formula for two elemental deep feature learning paradigms, i.e., learning with class-level and pair-wise labels, formulated as follows:

$$L = log\left[1 + \sum_{j=1}^{L} exp\left(\gamma\alpha_n^j(s_n^j - \Delta_n)\right)\sum_{i=1}^{K} exp\left(-\gamma\alpha_p^i(s_p^i - \Delta_p)\right)\right], \begin{cases}\alpha_p^i = [O_p - s_p^i]_+,\\ \alpha_n^j = [s_n^j - O_n]_+,\end{cases} \tag{5.2}$$

in which n and p annotations in a pairwise label paradigm refer to dissimilar and similar pairs. γ is the scale factor. L and K are the number of dissimilar and similar samples. Δ_n and Δ_p are dissimilar and similar margins, and α_n^j and α_p^i are non-negative weighting factors for pairwise distances between dissimilar samples (i.e., s_n) and similar samples (i.e., s_p) respectively. In their calculation, $[\mathring{u}]_+$ is the "cut-off at zero" operation to ensure α_p^i and α_n^j are non-negative. O_p and O_n is the optimum similarity score for s_p and s_n respectively. To reduce the number of hyperparameters, $O_p = 1+m$, $O_n = -m$, $\Delta_p = 1-m$, and $\Delta_n = m$. Hence, there are only two hyper-parameters, i.e., the scale factor γ and the relaxation margin m. In a class-level paradigm, the dissimilar and similar terminologies are replaced with between-class and within-class terms using the same equation as Equation 5.2. Analytically, Circle Loss offers a more flexible optimization approach towards a more definite convergence target.

In our proposed model, Circle Loss is adopted for the first time for TRL in a triplet learning paradigm using a Triplet-based Siamese network. We collect an equal number of dissimilar and similar pairs in each batch (i.e., $L = K$, equal to a pre-defined batch size) to create a triplet of anchor, positive, and negative samples. This also creates a balanced batch (with an equal number of positive and negative samples) in each epoch to prevent overfitting and bias towards a particular class. Moreover, we adopted the "Hard" triplet strategy [202]. In comparison with the random triplet strategy, where triplets are selected randomly, applying this strategy disregards triplets that are too easy, thereby reducing the risk of overfitting the model.

5.5 Experiments

We evaluate SiamCircle on five real trajectory datasets using nine evaluation metrics in downstream task using three classes of metrics: trajectory ranking, trajectory similarity, and clustering. The following sections describe the experimental settings, datasets, baselines, and evaluation metrics.

Experimental Settings. SiamCircle is implemented in Tensorflow.[1] In the data augmentation, jumps are only applied randomly to $k \in [2, 5]$ number of samples. The additive noise is $d_s \in [0.5, 1]$, and the jump scale is $d_s \in [15, 20]$. In the convolutional layer, the filter size is 64, and the kernel size is (5, 5). The average pooling size is (2,2), and the dropout rate is 0.2. The dimension of the embedding space is $d = 10$. In Circle Loss, we set γ=8 and m=0.85. We report the average and standard deviation of results across ten runs for each experiment. The Wilcoxon signed rank test [208] has been applied to investigate the significant differences between the top two performing models and to rank algorithms based on their performances in different metrics. The SiamCircle is optimized with Adam Optimizer. The learning rate is initialized to 0.001 and decayed by half after every 5 consecutive epochs with no improvement in the loss. The maximum number of training epochs is set to 1000 with an early stop after 10 consecutive epochs without improvements in the loss. In the following sections, we present the details of datasets, baselines, and evaluation metrics used in our study.

Datasets. Five pedestrian datasets —*eth*, *hotel* [209], and *zara01*, *zara02*, and *students03* [210]—are utilized in the experiments. These are widely recognized benchmarks for group detection tasks using spatio-temporal data [37]. They contain the location and velocity of movements across multiple timeframes, including ground truth information on group membership. All five datasets are captured in a free setting. Table 5.1 shows the main characteristics of these datasets. Since the number of samples per dataset is limited, we adopted our novel data augmentation process described in Section 5.4.1 to generate augmented trajectories per class (or group). The augmented data is only used during the training process. Hence, we chose a higher ratio for the test dataset to create enough samples for our evaluation experiments. Specifically, first, we split the original data set into a train and test set (i.e., train:test, 40:60). Next, via our proposed data augmentation process, we further enriched the training dataset by generating a larger sample size (over 5000 samples) with higher variability, i.e., different distortion and jumps scales. The generated augmented trajectory dataset is used only in the training process, and the remaining original dataset has been used for evaluation.

Baseline. We compared SiamCircle with three other deep learning models, namely S-BiLSTM [201], TrajCL [193], and Trajectory2vec [194]. We use the released code and default parameters for all baseline methods except S-BiLSTM, which has no released code. We implement this method following their original study [201]. In

[1] The code is available at `https://anonymous.4open.science/r/SiamCircle-Trajectory-Representation-Learning-17C3/`

Table 5.1: The characteristics of Opentraj datasets. The columns indicate the name of the dataset, the captured area, and the data size for Train, augmented train (i.e., Train($\overline{T}$)) and Test splits in the form of *(Sample Size, Number of groups).

Datasets	Area	Train (T)	Train($\overline{T}$)	Test
Eth	224.704	(40, 15)	(5098, 15)	(96, 37)
Hotel	67.210	(12, 6)	(5027, 6)	(31, 15)
Zara01	171.504	(34, 16)	(5050, 16)	(80, 35)
Zara02	158.063	(71, 35)	(5164, 35)	(167, 82)
Student03	242.884	(185, 61)	(5242, 61)	(433, 194)

TrajCL, the cell size is set to 5 as it performed better than the default value, i.e., cell size = 100, as it was designed for city-scale datasets.

Evaluation Metrics. We have adopted nine evaluation metrics in downstream task using three classes of metrics. Specifically, the trajectory ranking task is evaluated using the top-k hitting ratio (k = 1, 5, 10). The trajectory similarity task is assessed using the *most similar search (MSS)* and *cross-similarity (CS)*. Finally, the trajectory clustering task is evaluated using two supervised metrics, i.e., Normalized Mutual Information (NMI) [211] and Fowlkes Mallows (FM) [212], and two unsupervised metrics, i.e., Davies Bouldin (DB) [127] and Silhouette (Si) [123]. The summary of these evaluation metrics is presented as follows:

Top-K Hitting Ratio. This metric examines the overlap of the *top-k*, for $k = 1, 5$, and 10, k, sorted distances between embeddings and the ground truth. Specifically, we use the Euclidean measure to calculate the distances between the original trajectory T and other trajectories in the test set. Analogously, we calculate the Euclidean distances between $E(T)$ and the embeddings of other trajectories in the test set. We then sort the obtained distances in both sets and find the number of common indexes across the window size of k. The higher overlap ratio shows better effectiveness of the measure.

Most Similar Search (MSS) [195]. This metric measures the quality of self-similarity among trajectories. Specifically, we randomly select two trajectories per class and store them in two separate subsets. Thus, each subset's size equals the number of classes (or groups) in each test dataset. Then, we calculate the pairwise distances between the embeddings of all trajectories in these two subsets and sort the obtained distances in ascending order per class. Ideally, the two trajectories from the same class should rank first among others. We calculated the average rank of the same-class trajectories across all classes and reported it as the MSS score. The lower score shows a better performance.

Cross-Similarity (CS) [195]. This metric evaluates the cross-similarity performance by measuring the extent to which a similarity measure maintains the distance between two trajectories from different classes, i.e., cross-similarity. CS score is formulated in Equation 5.3:

$$CS = \frac{|d(T_a, T_b) - d(\overline{T}_a, \overline{T}_b)|}{d(T_a, T_b)} \tag{5.3}$$

where T_a and T_b are the embeddings of two distinct original trajectories from the test dataset. $\overline{T}_a$ and $\overline{T}_b$ are the embeddings of their distorted counterparts respectively. $d(.,.)$ denotes the Euclidean distance between the two given trajectories. A smaller CS measure shows better performance.

Clustering performance. Retrieving the embedding of raw trajectories allows a clustering algorithm to better identify groups of similar trajectories. To test this, we selected two supervised evaluation metrics, namely NMI [211] and FM score [212], and two unsupervised metrics, i.e., DB score [127] and Si [123], to evaluate the performance of a K-mean Clustering Algorithm [213]. The number of ground truth groups, as indicated in Table 5.1, defines the number of clusters in the K-mean Algorithm. In summary, the NMI score measures the agreement of the two assignments, with the highest score being 1. The FM score is the geometric mean of the pairwise precision and recall, rating from 0 to 1. The Si score indicates how well clusters are separated from each other, ranging between -1 and 1. DB score is the average similarity measure of each cluster with its most similar cluster, where the minimum score is zero, with lower values indicating better clustering. In all other metrics, the higher score indicates better-defined clusters.

5.6 Results

In this section, the experiment's results are presented in two studies: (i) an ablation study to explore different design options and (ii) a comparative study to compare the performance of our proposed model against three baselines.

Ablation Study. Our ablation study includes component analysis to investigate the impact of different neural network components and loss function analysis to investigate the impact of different loss functions in our framework.

Component analysis: As described before, our model mainly includes one 2-D convolutional layer to extract spatio-temporal features in both dimensions of trajectories (i.e., $1 \times Conv2D$). Through this experiment, we compare the performance of $1 \times Conv2D$ with three other alternatives. Since in the literature, recurrent neural networks have been suggested for analyzing time-series data, we have explored the use of LSTM and GRU layers instead of convolutional layers in models *LSTM*, and *GRU*. Moreover, we explored the performance of two 2-dimensional convolutional layers (instead of one) in $2 \times Conv2D$. As shown in Table 5.2, the model $1 \times Conv2D$, which includes one layer of 2-dimensional convolutional layer, on average, ranked higher than all the other models in all classes of metrics across all datasets. This demonstrates that the $1 \times Conv2D$ component captures spatial and temporal dependencies in the data. Hence, this model is used in the following sections to conduct further experiments.

Loss function analysis. After selecting the best-performing model in terms of neural network design, in this section, we investigate the impact of different loss functions on the performance of the selected model. Thus, we trained our proposed neural network design, i.e., $1 \times Conv2D$, with Triplet loss [202], Contrastive loss [214], USR loss [215], and Circle Loss. As shown in Table 5.3, the model $1 \times Conv2D$ with Circle Loss, on average, ranked higher than all the other models in almost all metrics across all datasets. There is only one exception on the unsupervised clustering task where $1 \times Conv2D$ with Circle Loss ranked last in average rank (M_R) and the rank of each unsupervised metric (i.e., DB and Si). A deeper analysis of the data shows that the worst performance of these two measures is in the $Hotel$ dataset, which includes the smallest area with the smallest sample size in the original datasets (See Table 5.1). This might explain the poor performance of these two metrics, as both indicate how well clusters are separated. This task imposes more challenges in small areas and makes the clusters less distinguishable. Yet, the other loss functions did not consistently outperform in either of these measures. Hence, $1 \times Conv2D$ with Circle Loss, i.e., SiamCircle, is selected and used in the following sections to conduct comparative experiments.

Comparative Study. In this section, we compare the performance of our proposed model against three baselines, namely S-BiLSTM, TrajCL, and Trajectory2vec, in three different classes of metrics, including trajectory ranking, trajectory similarity, and clustering, using the nine evaluation metrics that are explained in previous sections. The results are depicted in Table 5.4.

Trajectory Ranking. SiamCircle performed significantly higher than the baselines in the top Hit-k measure in almost all cases except in the Hit-1 metric of the Hotel dataset. However, the highest performance (obtained by S-BiLSTM) is not statistically significant. Moreover, the SiamCircle outperformed with an average performance gap of 19% compared with the second best-performing model, i.e., BiLSTM. This shows that the embedding vectors retrieved by our model are helpful for trajectory ranking tasks to estimate the top similar trajectories.

Trajectory Similarity. In the Trajectory Similarity task, none of the models con-

Table 5.2: The result of the ablation study in component analysis. The highest ranks in each evaluation metric are shown in Boldface. R denotes the performance ranking. The M_R shows the average rank per category. Categories, separated by horizontal lines, include trajectory ranking, trajectory similarity, supervised clustering, and unsupervised clustering.

Loss Function	Metric	Datasets					R	M_R
		ETH	Hotel	Student03	Zara01	Zara02		
LSTM	Hit-1	0.792	0.581	0.703	0.7	0.711	2.5	2.8
	Hit-5	0.773	0.884	0.706	0.798	0.901	2.8	
	Hit-10	0.807	0.919	0.699	0.839	0.874	3.2	
	MSS	2.694	1.867	2.977	2.912	2.78	3.4	3.0
	CS	0.617	2.968	1.389	3.388	2.206	2.6	
	NMI	0.85	0.847	0.911	0.785	0.841	2.4	2.4
	FM	0.361	0.423	0.392	0.165	0.091	2.4	
	DB	0.685	0.602	0.54	0.607	0.313	**2.2**	2.7
	Si	0.322	0.254	0.32	0.238	0.368	3.2	
GRU	Hit-1	0.76	0.452	0.738	0.738	0.725	**2.2**	2.6
	Hit-5	0.812	0.89	0.726	0.775	0.889	2.5	
	Hit-10	0.848	0.932	0.743	0.811	0.866	3.0	
	MSS	2.528	1.6	2.754	3.441	2.712	2.4	2.6
	CS	0.9	2.983	1.34	3.988	1.906	2.8	
	NMI	0.828	0.83	0.902	0.79	0.836	3.2	3.2
	FM	0.307	0.37	0.355	0.179	0.074	3.2	
	DB	0.726	0.613	0.517	0.573	0.323	**2.2**	2.3
	Si	0.326	0.285	0.305	0.243	0.377	2.4	
2×Conv2D	Hit-1	0.625	0.613	0.645	0.712	0.711	3.1	2.8
	Hit-5	0.792	0.89	0.741	0.767	0.858	2.7	
	Hit-10	0.865	0.961	0.742	0.822	0.824	2.6	
	MSS	2.222	1.8	2.6	3.676	2.932	2.8	2.4
	CS	1.542	2.495	0.907	6.217	1.119	**2.0**	
	NMI	0.83	0.77	0.923	0.785	0.832	3.1	3.1
	FM	0.308	0.198	0.464	0.146	0.05	3.2	
	DB	0.708	0.6	0.552	0.625	0.325	3.2	2.7
	Si	0.325	0.218	0.322	0.254	0.393	2.3	
1×Conv2D	Hit-1	0.771	0.523	0.686	0.724	0.732	**2.2**	**1.8**
	Hit-5	0.84	0.88	0.738	0.791	0.909	**2.0**	
	Hit-10	0.892	0.953	0.744	0.843	0.877	**1.2**	
	MSS	2.211	1.56	2.687	2.806	2.717	**1.4**	**2.0**
	CS	2.229	2.864	1.039	5.172	1.502	2.6	
	NMI	0.858	0.878	0.911	0.811	0.849	**1.3**	**1.2**
	FM	0.39	0.517	0.399	0.223	0.102	**1.2**	
	DB	0.627	0.671	0.526	0.589	0.324	2.4	**2.2**
	Si	0.328	0.218	0.331	0.287	0.366	**2.1**	

Table 5.3: The result of the ablation study in loss function analysis. The highest ranks in each evaluation metric are shown in Boldface. R denotes the performance ranking. The M_R shows the average rank per category. Categories, separated by horizontal lines, include trajectory ranking, trajectory similarity, supervised clustering, and unsupervised clustering.

Loss Function	Metric	Datasets					R	M_R
		ETH	Hotel	Student03	Zara01	Zara02		
Triplet loss	Hit-1	0.76	0.677	0.479	0.6	0.669	1.8	1.9
	Hit-5	0.79	0.852	0.67	0.758	0.801	2.0	
	Hit-10	0.857	0.935	0.69	0.814	0.815	2.0	
	MSS	1.972	1.4	3.9	3.147	2.949	1.8	1.8
	CS	1.9	2.762	1.968	6.197	1.13	**1.8**	
	NMI	0.869	0.934	0.864	0.774	0.836	2.4	2.3
	FM	0.406	0.765	0.174	0.123	0.063	2.2	
	DB	0.669	0.588	0.504	0.521	0.288	2.7	2.7
	Si	0.345	0.434	0.322	0.286	0.334	2.8	
Contrastive loss	Hit-1	0.49	0.613	0.407	0.312	0.585	2.8	2.9
	Hit-5	0.623	0.761	0.544	0.503	0.754	3.0	
	Hit-10	0.749	0.848	0.587	0.585	0.77	3.0	
	MSS	2.389	1.733	3.046	3.912	3.712	2.8	2.8
	CS	2.316	10.926	0.972	21.021	1.028	2.8	
	NMI	0.79	0.845	0.87	0.788	0.844	2.6	2.6
	FM	0.227	0.436	0.212	0.164	0.07	2.6	
	DB	0.606	0.336	0.454	0.643	0.239	2.0	**2.0**
	Si	0.395	0.38	0.34	0.253	0.394	**2.0**	
USR loss	Hit-1	0.302	0.452	0.172	0.162	0.296	4.0	4.0
	Hit-5	0.331	0.742	0.193	0.268	0.337	4.0	
	Hit-10	0.424	0.806	0.229	0.35	0.352	4.0	
	MSS	3.722	1.8	7.085	5.912	7.305	4.0	3.7
	CS	2.254	3.315	2.289	8.919	1.934	3.4	
	NMI	0.801	0.82	0.858	0.751	0.838	3.6	3.7
	FM	0.248	0.34	0.152	0.056	0.049	3.8	
	DB	0.548	0.22	0.504	0.305	0.293	**1.7**	**2.0**
	Si	0.392	0.595	0.301	0.479	0.285	2.4	
Circle Loss	Hit-1	0.771	0.523	0.686	0.724	0.732	**1.4**	**1.1**
	Hit-5	0.84	0.88	0.738	0.791	0.909	**1.0**	
	Hit-10	0.892	0.953	0.744	0.843	0.877	**1.0**	
	MSS	2.211	1.56	2.687	2.806	2.717	**1.4**	**1.7**
	CS	2.229	2.864	1.039	5.172	1.502	2.0	
	NMI	0.858	0.878	0.911	0.811	0.849	**1.4**	**1.4**
	FM	0.39	0.517	0.399	0.223	0.102	**1.4**	
	DB	0.627	0.671	0.526	0.589	0.324	3.6	3.2
	Si	0.328	0.218	0.331	0.287	0.366	2.8	

sistently outperformed in either of the metrics, i.e., MSS and CS. Yet, SiamCircle ranked highest in the MSS score, and Trajectory2vec ranked highest in the CS score. This means that SiamCircle was more successful in identifying trajectories belonging to the same class but faced difficulty distinguishing trajectories from different

Table 5.4: The result of the comparative study in the format of (mean±standard division). The statistically significant results are shown by *. The highest ranks in each evaluation metric are shown in Boldface. R denotes the performance ranking. The M_R shows the average rank per category. Categories, separated by horizontal lines, include trajectory ranking, trajectory similarity, supervised clustering, and unsupervised clustering.

Model	Metric	Datasets: ETH	Hotel	Student03	Zara01	Zara02	R	M_R
S-BiLSTM	Hit-1	0.449±0.497	0.561±0.496	0.42±0.494	0.34±0.474	0.554±0.497	1.8	1.9
	Hit-5	0.651±0.209	0.78±0.187	0.582±0.236	0.488±0.215	0.661±0.239	2.0	
	Hit-10	0.699±0.176	0.854±0.15	0.614±0.204	0.587±0.145	0.687±0.228	2.0	
	MSS	2.867±2.448	1.6±0.841	3.315±6.792	3.056±2.275	3.992±4.735	2.0	2.2
	CS	4.567±0.854	$0.594 \pm 0.127^*$	$0.635 \pm 0.148^*$	24.026±14.074	1.566±0.44	2.4	
	NMI	0.799±0.01	0.859±0.036	0.864±0.006	0.763±0.014	0.851±0.009	**1.8**	2.1
	FM	0.245±0.025	0.455±0.103	0.188±0.021	0.124±0.034	0.1±0.029	2.4	
	DB	0.474±0.023	0.391±0.041	0.466±0.023	0.39±0.048	0.281±0.021	1.6	1.9
	Si	0.415±0.014	0.4±0.05	0.31±0.012	0.323±0.032	0.388±0.019	2.2	
TrajCL	Hit-1	0.097±0.296	0.106±0.308	0.071±0.257	0.094±0.291	0.079±0.27	3.8	3.5
	Hit-5	0.194±0.188	0.476±0.323	0.182±0.209	0.204±0.179	0.192±0.216	3.4	
	Hit-10	0.288±0.19	0.566±0.167	0.234±0.172	0.32±0.167	0.276±0.181	3.2	
	MSS	4.153±6.816	2.87±1.56	21.125±34.181	9.103±10.186	5.36±8.493	3.8	3.4
	CS	4.966±1.819	27.072±17.447	1.976±1.789	3.236±1.477	1.356±0.243	3.0	
	NMI	0.731±0.01	0.754±0.015	0.836±0.005	0.756±0.008	0.782±0.002	4.0	3.8
	FM	0.158±0.024	0.171±0.055	0.136±0.014	0.126±0.024	0.068±0.002	3.6	
	DB	0.486±0.069	0.422±0.102	$0.405 \pm 0.024^*$	0.335±0.108	$0.073 \pm 0.022^*$	**1.4**	**1.2**
	Si	0.435±0.056	0.412±0.082	$0.436 \pm 0.048^*$	$0.438 \pm 0.066^*$	$0.518 \pm 0.007^*$	**1.0**	
Trajectory2vec	Hit-1	0.198±0.398	0.319±0.466	0.063±0.243	0.154±0.361	0.275±0.446	3.2	3.5
	Hit-5	0.134±0.142	0.401±0.272	0.096±0.131	0.23±0.237	0.376±0.304	3.6	
	Hit-10	0.16±0.107	0.455±0.141	0.117±0.11	0.268±0.135	0.385±0.291	3.8	
	MSS	5.572±5.282	1.953±1.179	15.701±16.687	5.624±6.107	4.766±4.223	3.2	2.6
	CS	1.187 ± 0.018	1.174±0.023	0.973±0.016	$1.424 \pm 0.014^*$	2.203±0.219	**2.0**	
	NMI	0.775±0.008	0.843±0.018	0.841±0.003	0.807±0.008	0.85±0.003	2.4	2.4
	FM	0.189±0.017	0.416±0.064	0.08±0.009	0.237±0.034	$0.149 \pm 0.01^*$	2.4	
	DB	0.638±0.028	0.472±0.028	0.611±0.013	0.564±0.035	0.31±0.014	3.2	3.5
	Si	0.277±0.011	0.388±0.028	0.221±0.006	0.262±0.014	0.324±0.009	3.8	
SiamCircle	Hit-1	$0.771 \pm 0.420^*$	0.523±0.449	$0.686 \pm 0.464^*$	$0.724 \pm 0.447^*$	$0.732 \pm 0.443^*$	**1.2**	**1.0**
	Hit-5	$0.840 \pm 0.144^*$	$0.880 \pm 0.118^*$	$0.738 \pm 0.195^*$	$0.791 \pm 0.146^*$	$0.909 \pm 0.116^*$	**1.0**	
	Hit-10	$0.892 \pm 0.086^*$	$0.953 \pm 0.055^*$	$0.744 \pm 0.155^*$	$0.843 \pm 0.1^*$	$0.877 \pm 0.146^*$	**1.0**	
	MSS	$2.11 \pm 2.434^*$	1.56±0.828	$2.687 \pm 3.702^*$	2.806±2.659	$2.717 \pm 2.358^*$	**1.0**	**1.8**
	CS	2.229±0.063	2.864±0.046	1.039±0.128	5.172±0.432	1.502±0.151	2.6	
	NMI	$0.849 \pm 0.009^*$	0.858±0.026	$0.921 \pm 0.004^*$	$0.805 \pm 0.007^*$	0.838±0.003	**1.8**	**1.7**
	FM	$0.366 \pm 0.025^*$	0.458±0.083	$0.451 \pm 0.019^*$	0.213±0.018	0.07±0.007	**1.6**	
	DB	0.668±0.025	0.618±0.046	0.529±0.015	0.584±0.023	0.328±0.012	3.8	3.4
	Si	0.329±0.012	0.257±0.024	0.331±0.007	0.289±0.011	0.379±0.006	3.0	

classes. Trajectory2vec uses an autoencoder with no explicit labeling to reconstruct representations. As a result, trajectory representation is solely based on the individual trajectory itself, without considering similarities with trajectories from the same class. This might explain why Trajectory2vec has performed better in distinguishing dissimilar trajectories (i.e., CS score) and did not perform at the same level as SiamCircle in identifying similarities among same-class trajectories (i.e., MSS score). Overall, SiamCircle obtained the best average rank M_R in the Trajectory Similarity task compared to the baselines.

Clustering. SiamCircle performed higher than the other three baselines in supervised clustering measures, i.e., NMI and FM scores (with S-BiLSTM performing similarly to SiamCircle in the NMI metric). However, this was not the case for unsupervised clustering measures, i.e., DB and Si. This could be due to providing explicit positive and negative samples for training in these two models, which allows the models to learn more specific patterns and relationships. Additionally, the dynamic penalty strategy by Circle Loss likely helped SiamCircle achieve a higher rank than S-BiLSTM. On the other hand, the TrajCL has ranked best in unsupervised clustering metrics. This could be because TrajCL did not use group membership during training. Thus, the obtained embeddings via TrajCL form clusters in the embedding space that do not necessarily reflect individual group memberships. This makes TrajCL particularly useful when ground truth data is not available. Yet, a deep understanding of data is required to translate the findings into an implication.

Overall, SiamCircle has ranked highest compared to the baselines in two classes of metrics, Trajectory Ranking and Trajectory Similarity. Meanwhile, in Trajectory Clustering, the highest performance is obtained only in Supervised metrics. This demonstrates the strong capability of representations obtained by SiamCircle to be used in various domains and applications.

5.7 Conclusion

This study revisits the problem of TRL, specifically when trajectories are collected from free movements in small areas. We introduced a novel framework, SiamCircle, which includes a Siamese framework with Circle Loss. We conducted experiments with five benchmark datasets, using nine evaluation metrics in downstream tasks using three classes of metrics, namely trajectory ranking, trajectory similarity, and clustering. In the ablation study, we demonstrated that one 2-dimensional convolutional layer together with Circle Loss, on average, outperformed other candidates. In a comparative study, SiamCircle consistently outperformed other models in trajectory ranking with an average performance gap of 19% compared with the second best-performing model, i.e., BiLSTM, and in unsupervised clustering measures. improvements over the existing measures. SiamCircle can be used in various tasks, such as analyzing social interactions in free settings, e.g., schoolyards and sports clubs. In future research, automating hyperparameter selection in Circle Loss based on the characteristics of the given datasets can be explored.

CHAPTER 6

Modeling Group Behaviour via Spatio-temporal Data

6

The contents of this chapter are based on the following publications (with permission from Springer Nature):

- M. Nasri, Z. Fang, M. Baratchi, G. Englebienne, S. Wang, A. Koutamanis, and C. Rieffe, "A gnn-based architecture for group detection from spatio-temporal trajectory data," in *International Symposium on Intelligent Data Analysis*, pp. 327–339, Springer, 2023 DOI: 10.1007/978-3-031-30047-9_26.

- M. Nasri, T. Maliappis, C. Rieffe, and M. Baratchi, "T-dante: Detecting group behaviour in spatio-temporal trajectories using context information," in *International Symposium on Intelligent Data Analysis*, pp. 28–39, Springer, 2024 DOI: 10.1007/978-3-031-58553-1_3.

Abstract

Modeling group behavior can be used in schoolyards to understand children's social behavior and their interactions with the social environment. Several studies have addressed the problem of identifying group behavior through modeling spatio-temporal trajectories. In this chapter, we revisit this problem by conducting two parallel studies using dyad-based models and context-based models. In this context, children are modeled as nodes in a graph. Such a graph representation indicates, for example, the social network of children in a class or playgroup. Our proposed dyad-based model, i.e., WavenetNRI, models interactions between each pair of nodes (i.e., dyadic nodes) using their spatio-temporal trajectories. Whereas our proposed context-based model, i.e., T-DANTE, includes the context information, i.e., it models interactions of dyadic nodes by additionally considering the spatio-temporal trajectories of surrounding nodes. We conducted our experiments using two collections of datasets, namely Opentraj datasets (with five real-world pedestrian datasets) and spring simulation dataset (with five simulation datasets), and two evaluation metrics, i.e., group mitre and group correctness. Our experiments compare the performance of these two models with two other baselines. The results demonstrate that including context information can improve the accuracy of group behavior modeling in Opentraj datasets. Meanwhile, in the simulation dataset, which includes groups with larger sizes, the dyad-based model performs better than other models.

6.1 Introduction

Children constantly interact with their social environment (i.e., peer groups and teachers) in schoolyards through different games and activities [1]. The social environment provides opportunities for children to develop their social skills. Yet, these environments may include barriers (e.g., environments in which children experience ostracism by their peers) that hinder social development for certain children. Addressing the existing barriers in the social environment is crucial to prevent problems such as bullying and social exclusion and promote emotional well-being in children [216].

The primary step in identifying these barriers is understanding children's social behavior and their group formations in schoolyards. The main challenge in obtaining this understanding is the variability of children's interactions over time and space. For instance, children often change their interacting partners or groups over recess time [35], or they often use several areas of schoolyards during group interactions [217]. Thus, accurately capturing children's group behavior in schoolyards is only possible when both spatial and temporal elements are considered in the design of the modeling framework [2].

Excessive literature in children's research has studied social behavior by measuring face-to-face interactions [218, 219]. In this form of interaction, children are physically present with one another within close proximity [77]. Similarly, in the context of the present thesis, proximity tags were initially adopted to capture children's face-to-face interactions in schoolyards (see Chapters 2 - 4). Although capturing face-to-face interactions provides an informative and straightforward measure to understand children's social behavior, it overlooks other forms of interactions that commonly happen in schoolyards. In order to create a clearer picture of children's social behavior, in addition to face-to-face interactions, we also captured parallel interactions, e.g., walking and running side-by-side, in schoolyards (see Chapter 5).

Yet, the complex nature of schoolyard activities might involve more complicated forms of interactions not captured in face-to-face contact or parallel movements. Understanding these complex group interactions requires analyzing children's behavior in their social networks in schoolyards. From the data science perspective, analyzing group behavior in a social network can turn into a mathematical problem: how to identify sub-groups (or sub-graphs) in a given community (or graph). To this end, various studies in social network analysis focused on designing community detection algorithms that identify sub-communities based on pairwise relationships among individuals [220–222]. Despite their great performance in identifying static groups, these algorithms might not be able to identify groups in scenarios where group formation dynamically changes over time and space. For example, when a group of children is playing hide and seek, depending on their role, they might be involved in social interactions that can only be revealed by analyzing the spatio-temporal dynamics of children in forming a group, i.e., how children move in space over time compared with their peers. Including these spatio-temporal dynamics is essential to analyze social interactions and group formations in a higher resolution, going beyond face-to-face interaction and parallel play.

The first attempts to address this challenge focused on classical machine learning models, which incorporate a manual feature extraction process to find the most significant features [223, 224]. Although the results show promise, the manual feature extraction and selection process is often time-consuming and might potentially introduce bias to the model. Recent studies in the field of artificial intelligence focused on developing neural network models to automatically extract features and identify sub-groups based on individual interaction graphs [225–228]. These pipelines typically incorporate spatio-temporal data to train a neural network model and reconstruct an affinity graph, i.e., a graph that represents the pairwise relationship of individuals. Applying a community detection algorithm or clustering method can identify sub-communities within this reconstructed graph. Previous studies focused specifically on spatial features by adopting multi-layer perception MLP layers [225] or 1-dimensional convolutional layers [226] to model group behavior, overlooking the temporal dependencies that might contribute to modeling group formations.

To address this gap, the current chapter revisits the problem of group behavior modeling in spatio-temporal data via neural network models by conducting two parallel studies: (1) WavenetNRI [3], built upon NRI [226], and (2) T-DANTE [4], building upon DANTE [225], to create an affinity graph via a neural network model that can be used by a community detection algorithm to identify groups in a given community. Overall, this chapter includes the following subjects:

- Discussing two novel neural network frameworks, i.e., WavenetNRI [3] and T-DANTE [4], to address group modeling tasks using spatio-temporal data.
- Discussing a trajectory simulation framework, built upon spring simulation framework [226, 229], to stimulate group and non-group interactions among particles in a physical system. This framework uniquely simulates attraction points (i.e., points where group members often mingle around) to stimulate group movements.
- Evaluating the performance of the two models, i.e., WavenetNRI and T-DANTE, using two sets of datasets, namely Opentraj dataset (includes five pedestrian datasets) and spring simulation dataset (with five simulation datasets) against two baselines (i.e., NRI [226], and DANTE [225]) via two evaluation metrics, i.e., Group Mitre and Group correctness.

The present study is organized as follows. The related literature is presented in Section 6.2. The group modeling problem is defined in Section 6.3. Section 6.4 presents the details of the proposed approach in pairwise information and context information. Section 6.5 presents the datasets, the evaluation metrics, baselines, and implementation details adopted in our experiments. Moreover, this section discusses the results of our experiments. Finally, Section 6.6 summarizes the study and points out the limitations and directions for future research.

6.2 Related Work

The spatio-temporal data adopted in modeling group behavior can be categorized into two areas: **Dyad-based modeling** and **Context-based modeling**. In dyad-based modeling studies, the spatio-temporal data per pair of nodes in the interaction graph focuses on training their model and predicting the affinity score. Studies in context-based modeling included the spatio-temporal data of the surrounding nodes, i.e., social context, in addition to the pairwise interaction data to predict the affinity score. The following section discusses the existing literature on adopting these two strategies.

6.2.1 Dyad-based Modeling

Various studies in this area adopted graph-based neural networks (GNN) to estimate pairwise interactions among agents [226–228]. Thompson et al. [230] modeled a scene as an interaction graph, where nodes and edges represent individuals and their dyadic relationship, respectively. GNN is used afterward to predict the pairwise affinity, indicating the likelihood of pairwise interactions. In another attempt, Kipf et al. [226] proposed Neural Relational Inference (NRI), which predicts interactions between moving particles using spatio-temporal data. Both studies assumed interactions among specific pairs of agents remain constant throughout the entire timeframe. Yet, in real-world social settings, individuals often change their interaction partners.

Moreover, they both overlook the symmetric group relationships among pairs in the affinity graph. Implementing this feature satisfies the following condition in the embedding space (where the affinity graph is reconstructed): if A is in a group with B, B is also in the same group with A, to account for bidirectional relations in group memberships. The dyad-based model, WavenetNRI, adopts the dilated residual causal convolutional (GD-RCC) block [231] to capture short and long dependencies in spatio-temporal dynamics. Moreover, it uses symmetric temporal edge features and a symmetric edge updating process to address the symmetric property of group relationships.

6.2.2 Context-based Modeling

This line of research incorporates the context information, i.e., the surrounding agents, in addition to the dyad information, into the model's design. The underlying reason is that determining whether two individuals belong to the same group does not solely depend on the behavior of those two individuals. Additionally, the behavior of surrounding individuals could also impact this determination, e.g., in the context of the schoolyard, identifying two children running together as a group will be easier by including the fact that other children are playing in the sandpit as it suggests significant spatio-temporal differences between the two groups. In line with this idea, Swofford et al. [225] introduced DANTE, a neural network model that incorporates MLP layers. DANTE adopts the information of surrounding agents in addition to the pairwise interactions to learn graph representation for a single-frame scene. In another attempt, Lu et al. [232] introduced VGDTN, which adopts 1-dimensional convolutional layers to identify group behavior using both dyad and context information in a single-frame scene. Yet, the focus of both studies on a single-frame scene might overlook the temporal features that occur over multiple timeframes. Moreover, MLP models and 1-dimensional convolutionals are known to be incapable of effectively modeling time-series data such as spatio-temporal trajectories compared with recurrent neural network (RNN) models. To address this, the context-based model, T-DANTE, enhances the context information by including scenes with multiple timeframes instead of a single scene. Moreover, by implementing RNN layers, T-DANTE aims to capture short and long dependencies in the spatio-temporal data.

6.3 Problem Formulation

Consider a dataset D that includes the movement trajectories of M agents. Each movement trajectory $X_m = \{x_1, ..., x_t, ..., x_T\}$ indicates a consecutive sequence of spatio-temporal features x_t of agent m, $m \in (1, M)$ over a timeframe with T time steps, $t \in (1, T)$. Each dyadic agent may interact with the others over the given timeframe.

Firstly, we are interested in estimating the pair-wise relationships between dyadic agents by learning the affinity score $h^2_{(i,j)}$ between all dyadic agents i and j and assembling all scores to form an affinity graph. Secondly, we are interested in identifying groups $C = \{c_j | j \in [1, K]\}$ of agents in the created affinity graph ($1 \leq K \leq N$ is the number of groups) under three main assumptions: (1) the group relationships are constant in a time window, while agents could interact with other agents from a different group. (2) Agents of the same group share similar spatial behavior over a timeframe. (3) The size of the timeframe is fixed across the measurements.

The present chapter discusses two approaches to address this problem: (1) the dyad-based modeling and (2) the context-based modeling. The dyad-based model learns interactions between dyadic nodes using their spatio-temporal trajectories. Meanwhile, the context-based model learns dyadic node interactions by considering surrounding nodes' spatio-temporal trajectories.

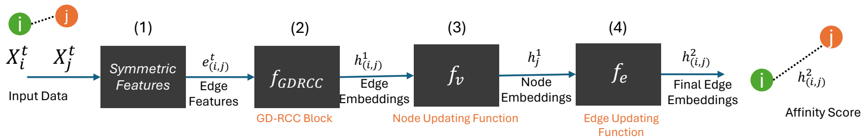


Figure 6.1: An overview of WavenetNRI framework. (1) The symmetric edge feature will be created based on the spatio-temporal of nodes i and j. (2) The edge embeddings will be created by applying a GD-RCC block to the symmetric edge feature sequences. (3) All edge embeddings will be aggregated per node j to obtain the node representation. (4) The node embedding and edge embeddings will be used in this block to create the final edge embeddings, i.e., affinity score, between node i and j.

6.4 Methodology

This section presents the design of two models, i.e., WavenetNRI and T-DANTE. These models identify group behavior by learning the affinity graph from spatio-temporal data of nodes (nodes can be agents or individuals depending on the context). While WavenetNRI focuses on extracting complex spatio-temporal dependencies based on dyad information, T-DANTE adopts an RNN model to identify group behavior using both dyad and context information. The details of these two models are described as follows:

6.4.1 WavenetNRI Framework

The WavenetNRI models group behavior solely based on dyad information. The design of this model is inspired by NRI framework [226] and Wavenet framework [231]. To satisfy the symmetric feature of group membership, WavenetNRI implements symmetric edge features and symmetric edge updating functions to account for the bidirectional nature of group memberships. Moreover, WavenetNRI adopts GD-RCC to learn short and long-term spatio-temporal dependencies in the edge feature. The overview of the WavenetNRI framework is depicted in Figure 6.1. This framework consists of four blocks, each representing one step in the training process as follows:

Step 1. Symmetric Edge Features. In the first step, the symmetric edge feature sequences $e^t_{(i,j)}$ will be created using the spatio-temporal data of dyad nodes i and j. The original NRI implements this by simply concatenating the spatio-temporal data of dyad nodes i and j (i.e., $e^t_{(i,j)} = [X^t_i, X^t_j]$). Built on this idea, WavenetNRI implements symmetric edge sequences to satisfy the symmetric nature of group relationships. The edge feature sequences in WavenetNRI are constructed by concatenating the pairwise distances and temporal increments in spatio-temporal data per dyad as follows:

$$e^t_{(i,j)} = [\|X^t_i - X^t_j\|, \Delta X^t_i \odot \Delta X^t_j], \quad t \in 1, ..., T-1, \quad \Delta X^t_i = X^{t+1}_i - X^t_i \tag{6.1}$$

Where $\|X^t_i - X^t_j\|$ denotes the Euclidean distance between dyad nodes i and j and $\Delta X^t_i \odot \Delta X^t_j$ denotes the element-wise production of the increments of dyads. Thus, edge feature sequence $e^t_{(i,j)}$ captures both spatial and temporal differences between each dyad. Moreover, the edge features are symmetric, i.e., $e^t_{(i,j)} = e^t_{(j,i)}$, corresponding to the symmetric properties of pairwise group relationships.

Step 2. GD-RCCC Block. The edge feature sequences $e^t_{(i,j)}$ obtained in the previous step will be given to the GD-RCC block to extract spatio-temporal features of the given edge (i.e., edge embeddings). The original NRI adopts one convolutional layer that may not efficiently capture the long-term interactions of edge feature sequences. In the WavenetNRI, a GD-RCC block [231] inspired by Wavenet [231] is used to transform the edge feature sequences $e^t_{(i,j)}$ into the edge embedding $h^1_{(i,j)}$ as formulated in Equation 6.2. The use of GD-RCC block has several advantages for the design of WavenetNRI: (1) The causal convolution maintains the order of the timely ordered edge sequences, (2) the dilated convolutional kernels exponentially expand the receptive fields, (3) the skip connection (i.e., 1D CNN) tackles the gradient vanishing problem, and (4) the gating activation function regulates the information flow.

$$h^1_{(i,j)} = f_{GDRCC}(e^t_{(i,j)}) \tag{6.2}$$

Step 3. Node Updating Function. This step sums up all edge embeddings $h^1_{(i,j)}$ for dyadic nodes i and j and gives them to the node updating function f_v (as proposed in the original NRI). This function generates higher level node embedding h^1_i (or h^1_j) for dyadic nodes i and j as formulated in Equation 6.3.

$$h^1_j = f_v(\sum_{i \neq j} h^1_{(i,j)}) \tag{6.3}$$

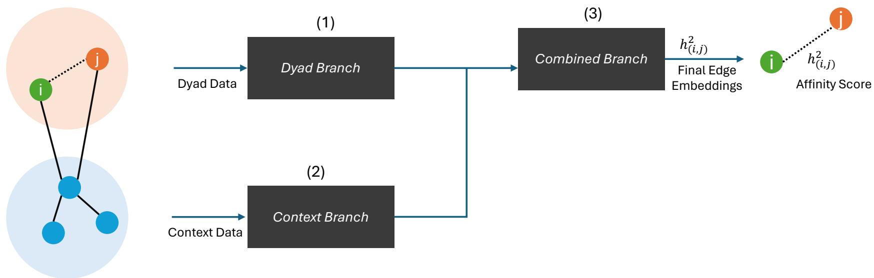


Figure 6.2: An overview of T-DANTE framework: The dyad branch extracts the spatio-temporal data of the pair of nodes of interest. The context branch extracts the spatio-temporal data of surrounding nodes, i.e., social context. The extracted features from the dyad branch and context branch will be merged in the combined branch. The output of this branch, i.e., the affinity score $h^2_{(i,j)}$, will be calculated per dyads to create the affinity graph.

Step 4. Symmetric Edge Updating Function. The element-wise production of the obtained node embeddings will be concatenated by the feature embeddings $h^1_{(i,j)}$ obtained in Step 2 and fed to the neural network f_e to get final edge embedding $h^2_{(i,j)}$ between dyadic node i and j, represented as the affinity score between the two nodes (See Equation 6.4). Through this process, the final affinity score $h^2_{(i,j)}$ captures interactions between dyadic nodes i and j and their interactions with other nodes [226].

$$h^2_{(i,j)} = f_e([h^1_{(i,j)}, h^1_i \odot h^1_j]) \tag{6.4}$$

During the supervised training phase, the ground-truth pairwise group relationships $G_{(i,j)}$ are used as labels. Since the datasets include an imbalanced distribution of the labels, the weighted cross-entropy is adopted as a loss function. This loss function assigns higher weights to the rare labels to compensate for their lower distribution. By minimizing the weighted cross-entropy, WavenetNRI is optimized to identify the "interaction" versus "no interaction" relation between nodes. After supervised training, the affinity graph will be constructed by assembling all the obtained affinity scores between pairs of nodes. The Louvain community detection algorithm [222] is applied afterward to the obtained affinity graph to find sub-groups among given nodes.

6.4.2 T-DANTE Framework

While the previous approach, WavenetNRI, solely focuses on the spatio-temporal data of the dyadic nodes, the T-DANTE framework additionally includes the spatio-temporal data of the surrounding nodes, i.e., social context. This section presents the details of T-DANTE, inspired by the DANTE framework [225]. T-DANTE extends DANTE by utilizing RNN blocks to retain temporal information and spatial features. During the training process, first, these spatio-temporal features will be extracted from the data to estimate the affinity score between pairs of nodes. Then, this affinity score will be compared with the pairwise group membership (i.e., ground truth $G_{(i,j)}$) using the log loss function. Assembling the estimated affinity score between all pairs of nodes creates an affinity graph, which can be used afterward in a community detection algorithm to detect sub-groups in the given data. The proposed T-DANTE consists of three branches: (1) Dyad Branch, (2) Context Branch, and (3) Combined Branch (See Figure 6.2). The dyad branch, similar to the WavenetNRI framework, captures local spatio-temporal information from the dyadic nodes. The context branch captures the spatio-temporal information from surrounding nodes (i.e., social context). The combined branch combines the output of these two branches and estimates the affinity score between the pair of nodes that are given in the dyad branch. The details of these branches are explained as follows:

6

1. Dyad Branch. The Dyad Branch extracts the spatio-temporal features of dyadic nodes using RNN layers, i.e., LSTM layers. The LSTM includes memory cells and gating mechanisms, to selectively store and retrieve information over long sequences, e.g., time series data such as movement trajectories. A series of convolutional layers are then applied to concatenate the spatio-temporal features extracted from RNN layers. Lastly, the dyad branch is followed by a Dropout layer to reduce overfitting and a Batch Normalisation layer to avoid the covariate shift and enhance the model's generalizability.

2. Context Branch. Context Branch follows the same identical design as the Dyad Branch. Yet, its given input data and role in the overall framework are different. The Context Branch extracts the spatio-temporal features of the surrounding nodes to account for context information. The number of surrounding nodes is a hyperparameter of the model (i.e., context size).

3. Combined Branch. The Combined Branch merges the extracted spatio-temporal features obtained from the Dyad Branch and Context Branch together.

Specifically, the extracted features are first flattened and passed through a series of fully connected layers, dropout layers, and batch normalization layers. Their specifications (e.g., number of layers, kernels, and filter size) depend on the characteristics of the dataset, such as the number of frames, the maximum number of nodes, and the data size. The last layer of this branch is a fully connected layer with a Sigmoid activation function to constrain the single output to the $[0, 1]$ range. This output is the affinity score for the dyadic nodes.

Assembling all the affinity values between all dyadic nodes creates the affinity graph. The group structures in the affinity graph will be identified afterward using the Dominant Sets (DS) community detection algorithm [233].

6.5 Experiments

We conducted several experiments to evaluate the models' performance. The following sections describe the datasets and evaluation metrics we used in our experiments. Furthermore, the baselines used to compare with the performance of the WavenetNRI and T-DANTE models are explained. Lastly, the implementation details of our experiments and the obtained results are presented. Our results compare the performance of the proposed frameworks, WavenetNRI and T-DANTE, with two other baselines using two evaluation metrics on two collections of datasets. The goal of this experiment is to address the following research questions:

6

- RQ. 1. Can symmetric edge features and GD-RCC block improve the performance of WavenetNRI compared with the original NRI model?
- RQ. 2. Can RNN block and including multiple timeframes per scene improve the performance of T-DANTE compared with the original DANTE model?
- RQ. 3. Which of the dyad-based or context-based models can perform better in identifying group behavior?

6.5.1 Datasets

In order to evaluate the performance of the models, we conducted our experiments using two publicly available datasets, i.e., Opentraj Dataset and Spring Simulation Dataset. Table 6.1 presents the characteristics of these datasets.

- **Opentraj Dataset.** Opentraj dataset [234][1] is extensively used in human trajectory prediction literature. This dataset includes the trajectories of pedestrians, location, and velocity in multiple timeframes, captured via static camera. Five pedestrian datasets, *eth*, *hotel* [209], and *zara01*, *zara02* and *students03* [210], which include the ground truth of the group membership, have been used in our experiments. This ground truth is created by annotating the pedestrians who seemed to walk in groups. The original dataset includes location data relative to the world reference W. In order to enhance the generalizability of our approach across different datasets, the trajectory of each pedestrian is transformed to a local coordinate system L_{ij}, defined as the middle point between pedestrian i and j.

- **Spring Simulation Dataset.** The spring simulation framework, built upon previous studies [226, 229], is developed to simulate group and non-group interactions among particles in a physical system. In line with the original studies, our spring simulation framework simulates the movements of groups of particles in a 2-D space. In their movements, those from the same group attract each other and distract from particles from another group. The locations, velocities, and the group membership (i.e., ground truth) of the particles are included in this simulation. Our proposed framework has made two improvements to the original framework: (1) defining group size as a simulation parameter that can be controlled over different experiments and (2) designing attraction points that stimulate particles from the same group toward certain pre-defined spots. In the proposed framework, pre-defined forces stimulate particles toward attraction points. All forces have the same strength, but their direction is different to point a particle towards a certain attraction point.

6

6.5.2 Evaluation Metrics

In order to evaluate the performance of the proposed model, we used two evaluation metrics, i.e., Group Mitre [235], Group Correctness [225, 233], in our experiments. In both evaluation metrics, due to the higher number of pairs from different groups compared to those from the same group, we adopted an F-1 score as it is more suitable for evaluating imbalanced datasets.

- **Group Mitre (G_M) [235]** is an evaluation metric, built upon the Mitre loss [236], has been used by several studies [229, 237, 238] to measure the

[1] https://github.com/crowdbotp/OpenTraj

Table 6.1: Characteristics of five Opentraj datasets and five spring simulation datasets used in both models, regarding the duration of measurements, pedestrian dataset in seconds and spring simulation dataset in timeframes, the number of agents, and the number of groups.

	Dataset	Duration	Agents#	Groups#
Opentraj	eth	773.4	360	58
	hotel	722.4	390	41
	zara01	360.4	148	45
	zara02	420.4	204	58
	students03	215.6	428	101
Simulation	sim_1	50	8	2
	sim_2	50	9	2
	sim_3	50	9	3
	sim_4	50	10	2
	sim_5	50	10	4

quality of the identified groups. Mitre loss adopts spanning trees to represent groups. This form of representation overlooks singletons, i.e., a group with only one node. Group Mitre solves this problem by adding a fake counterpart to each node. This fake node is considered in the same group as the original node only if the original node was singleton. The detailed implementation of G_M is presented by Solera et al. [235].

- **Group Correctness (G_c) [225, 233]** considers a group as correctly identified if at least $P * |c_j|$ of its members are correctly classified in the group, where $P \in [0, 1]$ is a threshold and $|c_j|$ indicates the size of the original group j. The $P = 1$ requires all agents in ground truth group membership data to be correctly identified in group j. Accordingly, $P < 1$ applies a milder metric in evaluating the quality of the identified groups.

6.5.3 Baselines

In the comparative study, we compared the performance of our proposed methods with two other baseline methods, namely NRI [226] as a dyad-based model and DANTE [225] as a context-based model. These baseline methods are described in Section 6.2 and are implemented based on their available source code. The original studies [4, 229] include extensive experiments where more baselines and dataset configurations have been presented. In this chapter, the most relevant approaches has been selected and presented in the result section for consistency and improving

Table 6.2: The results of Group Correctness G_C and Group Mitre G_M for WavenetNRI and T-DANTE compared with baselines using Opentraj datasets and simulation datasets. The ***** sign shows that this result is significantly different compared with other cases under the same evaluation metric and dataset.

	Pedestrain Dataset									
	eth		*hotel*		*zara01*		*zara02*		*students03*	
	G_C	G_M	G_C	G_M	G_C	G_M	G_C	G_M	G_C	G_M
DANTE	0.319 ±0.047	0.548 ±0.019	0.431 ±0.043	**0.586 ±0.035**	0.731 ±0.051	0.793 ±0.028	0.633 ±0.038	0.705 ±0.026	0.024 ±0.012	0.502 ±0.013
NRI	0.201 ±0.062	0.571 ±0.074	0.169 ±0.054	0.540 ±0.097	0.285 ±0.067	0.597 ±0.053	0.106 ±0.035	0.417 ±0.019	0.006 ±0.010	0.280 ±0.026
WavenetNRI	0.242 ±0.059	0.553 ±0.057	0.202 ±0.048	0.455 ±0.080	0.361 ±0.091	0.627 ±0.066	0.184 ±0.065	0.462 ±0.040	0.001 ±0.004	0.280 ±0.024
T-DANTE	**0.590* ±0.030**	**0.665 ±0.017**	**0.508* ±0.043**	0.542 ±0.023	**0.821* ±0.015**	**0.838* ±0.015**	**0.870* ±0.011**	**0.873* ±0.011**	**0.696* ±0.056**	**0.780* ±0.028**
	Simulation Dataset									
	sim_1		sim_2		sim_3		sim_4		sim_5	
	G_C	G_M	G_C	G_M	G_C	G_M	G_C	G_M	G_C	G_M
DANTE	0.215 ±0.007	0.717 ±0.004	0.198 ±0.008	0.701 ±0.003	0.095 ±0.011	0.518 ±0.011	0.199 ±0.011	0.712 ±0.005	0.041 ±0.007	0.425 ±0.009
NRI	0.984 ±0.004	0.991 ±0.002	0.983 ±0.007	0.993 ±0.002	**0.988* ±0.004**	**0.995* ±0.002**	0.996 ±0.003	0.999 ±0.001	**0.988* ±0.007**	**0.995 ±0.003**
WavenetNRI	**0.996 ±0.006**	**0.998 ±0.002**	**0.995* ±0.004**	**0.998* ±0.001**	0.977 ±0.008	0.988 ±0.004	**0.998* ±0.004**	**0.999 ±0.001**	0.953 ±0.011	0.968 ±0.009
T-DANTE	0.969 ±0.002	0.983 ±0.002	0.980 ±0.002	0.989 ±0.001	0.982 ±0.006	0.988 ±0.003	0.971 ±0.006	0.987 ±0.002	0.945 ±0.011	0.976 ±0.003

the readability.

6.5.4 Implementation Details

Our experiments were implemented via the Python programming language. The details of the implementation of WavenetNRI[2] and T-DANTE (with the spring simulation framework)[3] are available in the GitHub repository. We split each Opentraj dataset into 5 folds and evaluated the performance of each method across 5 times experiments per fold (i.e., 25 runs in total per method). Since spring simulation datasets were generated under controlled conditions, they have not been split into folds. Each spring simulation dataset was randomly split into train, test, and validation datasets. The performance of each method has been evaluated across 25 times experiments per simulation dataset. We investigated the significant differences between the top two performing models by implementing the Wilcoxon signed rank test [239]. In the following sections, the performance of the proposed models against three state-of-the-art baseline methods is presented.

[2] https://github.com/fatcatZF/WavenetNRI
[3] https://github.com/ADA-research/context-group-detection

6.5.5 Results

In this section, the performance of WavenetNRI and T-DANTE is compared with the baselines. The results of the experiments for the simulation dataset and Opentraj dataset are presented in Table 6.2.

6.5.5.1 Opentraj Datasets

According to Table 6.2, WavenetNRI has outperformed NRI in most of the cases across different Opentraj datasets. This answers RQ. 1, indicating that overall, including symmetric edge features and GD-RCC block to capture more complex dependencies in the spatio-temporal data has improved the performance of the WavenetNRI.

In order to answer RQ. 2, we compared the performance of DANTE with T-DANTE. Our results show that T-DANTE is the superior model using the Group Mitre metric in all datasets, except in the hotel dataset, in which DANTE performs better. Yet, this result is not statistically significant. The superiority of T-DANTE against DANTE in most cases demonstrates that including temporal dependencies via the LSTM layers and further enriching with multiple timeframes per scene has enhanced the performance of T-DANTE. Thus, implementing LSTM layers is more suitable compared with MLPs when using the Opentraj datasets.

Finally, RQ. 3 compares the two dyad-based models, i.e., NRI and WavenetNRI, with context-based models, i.e., DANTE and T-DANTE. The result shows that T-DANTE outperforms all baselines, i.e., DANTE, NRI, and WavenetNRI, for all Opentraj datasets using the Group Correctness metric. This shows that, indeed, including context data is beneficial for modeling group behavior.

6.5.5.2 Spring simulation datasets

According to Table 6.2, inline with the result of Opentraj datasets, the capability of WavenetNRI in learning symmetric edge features and capturing complex dependencies via GD-RCC has enhanced its performance compared with NRI, addressing RQ. 1).

To address RQ. 2, we compared the performance of DANTE with T-DANTE. The results show that similar to the findings in the Opentraj datasets, including the LSTM layers and multiple timeframes in T-DANTE, have significantly improved its performance compared with DANTE across different simulation datasets using both metrics.

Finally, to address RQ. 3, we compared the performance of dyad-based models, i.e., NRI and WavenetNRI, with context-based models, i.e., DANTE and T-DANTE.

The result of this comparison shows the superiority of dyad-based models over context-based models across all simulation datasets. This contrast with the result of Opentraj datasets can be explained by the differences in the characteristics of the Opentraj and simulation datasets. Compared with Opentraj datasets, spring simulation datasets have more scenes with group sizes larger than 3 particles. This feature makes it suitable for dyad-based models, i.e., NRI and Wavenet, to extract contextual information without being limited to the number of surrounding nodes.

Overall, the results in both dyad-based and context-based models demonstrate the positive impact of capturing temporal dynamics in the data, either with GD-RCC block or LSTM layers. Moreover, the context-based models were able to more accurately model datasets with smaller group sizes, which are mainly included in the Opentraj datasets. Whereas the dyad-based models were able to more effectively extract spatio-temporal features in larger group sizes that are mainly included in the simulation datasets. Additionally, in the Opentraj dataset, pedestrians often come to the scene from one of the two ends (of streets) and leave the scene from the other end. This structured movement might create overshared trajectories between different groups and pose challenges to modeling group behaviors. Thus, including context information has enhanced the performance of these models. In the spring simulation dataset, particles freely move in a physical box, and their movement is only directed by pre-defined attraction points. This is similar to scenarios where individuals have unstructured movements with relatively mild restrictions. In these types of scenarios, dyad information provided sufficient information to model group behaviors, and adding context information did not improve the performance. For example, on university campuses or on urban pavements where individuals appear in smaller group sizes with structured movements, context-based models can more accurately identify group behavior. Whereas in the context of individuals with unstructured movements in larger groups, such as athletes on a soccer field or children in schoolyards, dyad-based models can more accurately identify group behavior. Thus, either of these models might be useful for a specific scenario, depending on its characteristics. This leads us to the necessity of implementing dynamic context size in our model to automatically define context size based on the characteristics of the given dataset or even the specific scene.

6.6 Conclusion

Analyzing children's group behavior in schoolyards enables us to identify limitations and possibilities in social environments around the child. In dynamic social settings, such as children in schoolyards, individuals constantly change their interaction partners and activities, which poses extra challenges to modeling group behavior. In these scenarios, analyzing group behavior requires including both spatial and temporal elements in individuals' movements. To address this challenge, the present study aims at modeling group behavior using spatio-temporal data by conducting two parallel studies, dyad-based modeling and context-based modeling. The first study, i.e., WavenetNRI as a dyad-based model, is built up on NRI [226] and Wavenet [231] frameworks. WavenetNRI implements two features: (1) symmetric edge features with symmetric edge updating processes to account for the symmetric nature of group membership and (2) GD-RCC block to capture complex spatio-temporal dependencies in data. This model solely adopts spatio-temporal data between dyadic nodes to train the neural network model and reconstruct the affinity graph. The second study, i.e., T-DANTE as a context-based model, is built on the DANTE framework. T-DANTE adopts LSTM layers to estimate the affinity scores using the spatio-temporal data of the surrounding nodes, i.e., context information, in addition to the data of dyadic nodes. Moreover, this framework includes multiple timeframes per scene to enrich the context data.

Our comparative study against state-of-the-art baselines demonstrates that T-DANTE is the superior model for modeling group behavior using real-world Opentraj datasets. Whilst WavenetNRI outperformed other baselines in simulation datasets. The superiority of T-DANTE versus other dyad-based models, e.g., WavenetNRI and NRI, in Opentraj datasets shows that including context information has enhanced the performance of group behavior modeling in Opentraj datasets where group sizes are relatively small. Moreover, the superiority of T-DANTE over the original DANTE shows that including RNN layers can better capture spatio-temporal dependencies compared with MLP models. On the other hand, WavenetNRI has outperformed the baselines in spring simulation datasets where larger group sizes are available. The superiority of WavenetNRI over the original NRI shows that including the symmetric edge features and GD-RCC block can better capture the spatio-temporal dependencies for modeling group behavior in larger social settings via dyad information.

This finding shows that our proposed method is capable of modeling group behavior using spatio-temporal data. Moreover, the design of our models is not limited to certain from of interactions, e.g., face-to-face interactions or parallel plays. This feature enables modeling complex group behavior in higher resolution. While in small social settings with structured movements, such as pedestrians' movement on pavements, including context information is beneficial for identifying group behavior, in larger social settings with relatively unstructured movements, such as children in schoolyards, focusing on dyad interactions is sufficient to model group behavior.

Due to the limited access to group membership data for children in schoolyards, we have only tested the performance of our proposed models on Opentraj benchmark datasets and spring simulating datasets and not on actual schoolyards. However, the Opentraj dataset has been collected from pedestrian movements in constrained environments, e.g., university campuses, which to some degree is comparable to schoolyard scenarios where children freely move in a constrained environment. Yet, applying our proposed method to children's datasets might require further investigation. For example, since children's group dynamics constantly change over time. Thus, the implementation of dynamic context size and dynamic group membership might be required in the design of the models.

Future research can explore the incorporation of dynamic context size (based on the presented number of nodes) and dynamic group membership per scene to enhance the generalization of the proposed approach across different datasets. Another future approach could be extending the proposed models in real-time applications with online data streaming. Various applications, such as analyzing students' social behavior in schoolyards, monitoring tourists' behaviors in touristic sights, and analyzing sports teams' performances, may benefit from the presented work.

CHAPTER 7

General Discussion and Conclusion

7.1 Summary

Playgrounds provide opportunities for children to grow and develop their social and physical skills. Yet, playgrounds might include barriers that hinder children's development, such as limited accessibility or discriminatory social behaviors. Playgrounds should become an environment where children feel safe, accepted, and included. Creating such an environment starts with understanding children's capacities and needs and then identifying playgrounds' limitations and possibilities in relation to these needs. While the availability of data has opened up new possibilities to tackle fundamental challenges in different communities, the question from a data science perspective is how to design an analysis framework that can effectively collect, process, and interpret data from children in playgrounds as objectively and comprehensively as possible and derive meaningful insights into playground dynamics. Designing such a framework enables us to address fundamental questions of various stakeholders, such as psychologists, designers, and policymakers.

The main goal of this thesis is to develop a data analysis framework for capturing children's behavior in playgrounds. The proposed framework addresses the three main characteristics of playgrounds that pose challenges to the design of our framework, i.e., multiple environments, individual experiences, and spatio-temporal dynamics, throughout Chapters 2 to Chapter 6. Specifically, our data-driven approach based on sensors, its feasibility to capture children's behavior comprehensively, consistently, and unobtrusively, and the identified interconnected environments are presented in Chapter 2. Chapter 3 proposes a novel spatio-temporal metric to measure the impact of physical designs, e.g., play structures, layout, etc., on children's spatio-temporal social networks. The proposed metric has also been used to analyze children's peer networks in playgrounds. Chapter 4 has examined the use of children's experiences in analyzing sensor data. The measure of loneliness via self-reports enabled us to explore how the duration of face-to-face contact in sensor data is correlated with the reported level of loneliness. In Chapter 5 and Chapter 6, we specifically focused on the spatio-temporal dynamics of individuals in social settings. Chapter 5 aims to model interactions beyond face-to-face contact by focusing on parallel movements to identify specific forms of interactions (e.g., bike riding, walking, or running side-by-side). Chapter 6 focuses on detecting interactions between pairs of individuals based on their movements to identify group behavior. This chapter further investigates the inclusion of movements of surrounding individuals, i.e., social context, and the interactions between pairs of individuals to model group behavior.

The following section presents the general discussion of these studies in greater detail.

7.2 General Discussion

By incorporating advanced data analysis techniques, our analysis framework addresses the three main characteristics of playgrounds that pose challenges to the design of our framework, i.e., multiple environments, individual experiences, and spatio-temporal dynamics. This section presents the main outcomes and general discussions per characteristics as follows:

7.2.1 Multiple Environments

In the context of multiple environments in playgrounds, we asked, **"RQ. 1. To what extent can modern sensing technologies capture individual interaction with multiple environments in playgrounds?"** To address this question, Chapter 2 adopted the theory of affordances [45, 240] to specify the environments around the child. Affordances are the actionable properties an environment presents to a child (e.g., a sand-pit affords to build a sandcastle) in relation to their desires, needs, and capacities, i.e., effectivities [45,240]. Accordingly, three types of affordances are identified in playgrounds, i.e., physical, social, and cultural. The physical affordances enabled us to include the impact of the physical design (e.g., play structures, layout, and materials in playgrounds). The social affordances examined children's social networks, while the cultural affordances defined the local rules and constraints (e.g., restrictions on using certain areas).

Moreover, in Chapter 2, we introduced our data-driven approach based on sensor data, which captures children's interactions with respect to these identified affordances. The sensor system adopted in this study includes GPS loggers, proximity tags, and MMR sensors to capture locations, face-to-face contacts, and physical activities, respectively. This sensor system enabled us to capture different aspects of children's behavior. For example, GPS loggers recorded the location of individuals in playgrounds, enabling us to understand which areas they used, how often, and for how long. Thus, GPS data mainly contributed to capturing the physical affordances.

On the other hand, the proximity tags captured children's face-to-face contact with peer groups and supervisors in the playground, addressing the social affordances around the child. Lastly, obtaining the physical activity level from MMR data and GPS and proximity data enabled us to verify the local rules and constraints in the cultural affordances. In addition to capturing multiple aspects, sensor data enabled us to include all participants simultaneously and unobtrusively over the entire break, whilst this was not possible using classical behavioral methodologies such as field observation.

These findings may benefit psychologists, designers, policy-makers, and relevant authorities. For instance, by providing an ethical and privacy-preserved approach to capture children's behavior via sensing technologies, schools can understand popular play structures and define a timetable for their use if necessary, especially if the crowd or noise is overwhelming for some children. They can also examine if certain rules or constraints have been followed by students, for example, when the use of a certain area is prohibited. Adopting our data-driven approach methodology also enables policymakers to validate specific playground policies. The primary steps have already been taken in this field by studying the role of movement data in facilitating evidence-based policy-making in schools, the summary of which is presented in the Data for Policy Conference [6].

The second question in this context was **"RQ. 2. To what extent does the physical environment around the child impact the child's social behavior in playgrounds during recess?"**. Chapter 3 focused on the impact of physical and social environments around the child. Specifically, this chapter introduced a novel spatio-temporal metric to examine the accessibility of individuals in a spatio-temporal social network. In a case study, this spatio-temporal accessibility metric has been adopted to analyze the spatio-temporal social network of children in two playgrounds. The results showed that, in general, children were more accessible to peers from the same group. Similarly, children more easily accessed peers from the same group. Moreover, the most accessible sub-groups utilized common areas, mainly around play structures during recess.

Schools, municipalities, designers, policymakers, and relevant stakeholders can benefit from this study by examining the spatio-temporal accessibility of children and analyzing the impact of social and physical affordances in playgrounds in different aspects.

First, municipalities, schools, and designers can optimize the layout of design elements to enhance accessibility in playgrounds. This includes strategically placing play structures (e.g., seesaw and sandpit) and physical features (e.g., bench and storage) so that children are equally accessible in their social network and can access their peer group equally. In this context, our proposed metric can validate a specific design regarding spatio-temporal accessibility. Similarly, policymakers can adopt our spatio-temporal metric to validate how implementing specific policies impacts children's accessibility in their social network.

Second, designers can use our proposed metric to validate whether their design caters to all children with different effectivities. Specifically, play structures might have different levels of impact on children's social and cultural affordances. Some require to define certain rules and constraints, e.g., areas for playing soccer with designated goals and borders, and some are more relaxed regarding rules, e.g., bench. Some design features might offer solid play, e.g., bar fix, while some might introduce collaborative games, e.g., sandpit or seesaw. Including these different designs with different affordances enables equal access to playground facilities and peer groups for children with different effectivities.

Third, analyzing and understanding this impact enables designers to implement line marks and designated areas for different purposes, e.g., pick-up and drop-off zones, to enhance safety in playgrounds.

7.2.2 Individual Experiences

While sensors provided continuous measurement from individuals, our subjective measurements provided more context, depth, and meaning to the sensor data. This has been the focus of Chapter 4, in which we used sensor measurements, self-reports, and peer nomination to understand children's experiences in playgrounds. In this context, we asked **"RQ. 3. To what extent does including individual differences in experiencing playgrounds facilitate a more accurate interpretation of data?"** Chapter 4 examined the social connectedness of autistic children in playgrounds compared to their non-autistic peers based on differences in effectivities [240]. We adopted various methods, including sensor data capturing children's face-to-face contacts, peer nomination data measuring their pairwise friendship status, and loneliness questionnaires to assess their feelings of loneliness. Our findings showed that according to peer nomination data, autistic children in this study had fewer reciprocated friendships than their non-autistic counterparts. Nevertheless, according to sensor data, both autistic and non-autistic children showed similar levels of peer interactions and similar levels of centrality in their social network during recess. Moreover, our self-report results showed no differences in the levels of loneliness in these two groups. These data combined suggested that the underlying reason for experiencing loneliness differed in the two groups of children: non-autistic children tended to feel lonelier (self-report) when they had shorter interaction time with peers during recess (sensor data), whereas autistic children reported higher loneliness levels (sensor data) when they were less well included as preferred playmates within their peer networks (peer nomination).

These outcomes show the importance of addressing differences in effectivities either at the group level or at the individual level in relation to the affordances of the environment. Accordingly, in our playground context, the rules and constraints (i.e., cultural affordances), physical features, accessibility, and capacity of sub-areas (i.e., physical affordances), group sizes, and peer populations (i.e., social affordances) are all essential elements in understanding individual differences and accommodating their effectivities. Therefore, it is crucial to consider different affordances in playgrounds, as well as differences in the effectivities of children in their interaction with these affordances. It is not about educating a child to fit in a community, but it is about adjusting the environment to accommodate the effectivities of all individuals.

This study highlights the necessity of multimodal methodology in our framework. Adopting modern sensing technology is an essential component for capturing children's behavior in the required level of detail. Sensors capture all participants' locations, activities, and interactions throughout every second of recess time. Yet, sensors alone cannot offer a clear picture of playground dynamics, as they struggle to accurately measure individual experiences, e.g., children's capacities and emotions (i.e., effectivities). For example, just because someone is alone in the playground does not imply loneliness; being with others does not mean feeling accepted and less lonely. Thus, extracting in-depth knowledge and insights from the data is only possible by combining modern sensing technologies with subjective measurements, such as self-reports, peer nominations, video observations, field observations, and questionnaires for parents and teachers, into the data collection methodology.

7.2.3 Spatio-temporal Dynamics

Analyzing and understanding children's social behavior plays a major role in examining their social affordances. In Chapters 2-4, social behavior has been defined by measuring face-to-face contacts. This is a known method adopted by various research studies in social and behavioral sciences [35,36,76]. Yet, social interactions may include more complex scenarios than face-to-face contact. In playgrounds, for example, children might bike together side-by-side, or they might play hide & seek. These types of play do not necessarily include face-to-face contact, yet they indicate strong forms of social interactions. To capture this, Chapter 5 and Chapter 6 particularly focused on analyzing complex spatio-temporal dynamics to obtain a clearer picture of social affordances around the child. We designed advanced machine learning algorithms via AI models to find meaningful patterns in the spatio-temporal data for detecting group interactions. Specifically, in Chapter 5, we asked, **"RQ. 4. To what extent can spatio-temporal data identify parallel movements as one specific form of social interaction?"**.

We addressed this question by designing a trajectory representation learning model, i.e., SiamCircle. The trajectory representations obtained via SiamCircle can be used to identify similar movements in a trajectory similarity computation task. While previous trajectory representation models predominantly focus on modeling structured movements within large-scale urban environments (e.g., vehicles or pedestrians on streets), this chapter tackles a more complex challenge of modeling free movements collected from small-scale social spaces, e.g., playgrounds. Our experiments demonstrate great promise in using SiamCircle to examine similar movement trajectories.

Yet, including more complex forms of group interactions embedded in the spatio-temporal dynamics of children makes our understanding of social environments around the child clearer. Therefore, we asked, **"RQ. 5. To what extent does spatio-temporal data of individuals enable us to model group interactions as one form of social behavior?"**.

Chapter 6 has addressed this question by discussing two parallel studies in detecting group interactions using individual spatio-temporal trajectories in constrained environments. The first study used WavenetNRI, a supervised learning approach based on graph neural networks for inferring complex group interactions from spatio-temporal data. WavenetNRI incorporates symmetric edge features and edge updating processes to account for symmetric group relationships. Moreover, WavenetNRI employs a gated dilated residual causal convolutional block to capture short and long dependencies in the spatio-temporal data. The second study proposed T-DANTE, which includes the contextual information embedded in the trajectories of the surrounding peers, along with the spatio-temporal dynamics of the individuals, to examine whether two pedestrians belong to the same interaction group. T-DANTE incorporates LSTM layers to capture the temporal dynamics inherent in the pedestrian's trajectories. In a comparative study, T-DANTE outperformed the other methods in the group detection task using real-world pedestrian datasets which include groups with smaller sizes. Whilst, WavenetNRI outperformed other models in simulation datasets, which include larger groups. This result shows that either of the models can be helpful in modeling group behavior depending on the characteristics of the given datasets.

This line of research provided valuable insights into complex social phenomena that occurred in the social affordances of micro-communities. In both chapters, the Opentraj benchmark dataset, due to the availability of ground truth information, has been used to train, test, and evaluate the proposed AI models. Since both Opentraj datasets and the playground datasets include free movements in constrained environments, the proposed AI models can potentially be used in playground settings to capture complex forms of group interactions among children. Understanding children's group behavior provides valuable insights for psychologists, designers, educational practices, and policy-makers to identify issues such as bullying and social isolation more effectively. Moreover, these insights enable schools and stakeholders to design and implement practical interventions to address the identified issues and create equity for all children.

Interdisciplinary collaboration is the key ingredient of this thesis. This project was primarily designed via a multidisciplinary collaboration and remained interdisciplinary in every step and decision. The choice of sensors in computer science, for example, depends on answering the question of 'what do we need to measure?'. From a psychological point of view, it is crucial to include social interactions and measures for analyzing children's social networks. From an architecture and design point of view, it is crucial to measure the use of space and the areas where children spend most of their time. Yet, there are also areas where psychology and architecture meet the same goal; for example, which locations offer more social interactions? These questions from different disciplines have formed the design of the sensors by computer scientists in our research. Yet, interdisciplinary collaboration is not only about choosing sensors and defining which variables to measure. We adopted excessive interdisciplinary knowledge to design our social network analysis and machine learning algorithms, reveal insights, and interpret the findings and patterns from sensor data. We required input from psychologists to understand the capacity and desires of children with autism. We needed the architect's input to understand the design concept and what environments offer to the users based on their capacities. We explored various playgrounds concerning the possibilities and limitations they might have for children in terms of the noise level, the density of use, the density of play structures, the variability of play the playground offers, etc., bearing in mind the effectivities of vulnerable children. All these different input sources are considered when designing a data analytics algorithm, e.g., social network analysis and machine learning model, tailored explicitly for analyzing children's behavior in playgrounds. Thus, interdisciplinary collaboration is essential in every step and every decision.

7.3 Limitations and possible directions for future works

This section discusses the limitations of the current work and points out the possible solutions to address the identified limitations in future work.

- **Enhancing the sensing system.** Our proposed sensing system accurately captured a wide range of subjects and activities. Yet, this system can be improved in several dimensions as follows:

  – *User-friendly hardware.* While younger children may find sensor-equipped belts engaging, these belts are unsuitable for older adolescents, for example, in high schools. The implementation of smartwatches or wristwearable devices can be investigated in future directions to enlarge the scope of the research. Moreover, future research may benefit from implementing a co-design process to choose the sensing technology that is most suitable for the target participants. In this process, the participants, including vulnerable children, help researchers to understand which design feature best suits their capacity and needs.

- *Accuracy and resolution.* GPS technology showed promise in tracking individual positions. However, it is unsuitable for indoor areas, e.g., classrooms or sports halls. Moreover, UWB technology (accuracy of 10 centimeters) might be a better fit for playground contexts than GPS (accuracy of 1–10 meters) regarding positional accuracy. Thus, future research may investigate using alternative technologies to enhance accuracy.
- *Real-time experiences.* Our current system lacks information about the value of contacts and movements in real time. For example, while we can detect a child playing alone, we do not know their underlying emotions or preferences at that specific moment. Future studies may study the inclusion of EMA (e.g., via smartwatches) to allow participants to express their emotions and preferences in real time.

- **Enriching datasets.** Our study only collected data from two special education schools, limiting the generalizability of our conclusions. One of the reasons was the COVID-19 crisis, which prevented conducting more data collection at schools. The other reason was that a limited number of parental consents were received due to privacy concerns. Future research may enrich the sample size in the following directions:
  - *Privacy-preserved data collection.* Implementing anonymous data collection aligned with GDRP and upon approval from the ethics committee might be a possible solution to maximize participation while preserving participants' privacy.
  - *Enhancing awareness about digital solutions.* The use of sensing technologies in children's research is a relatively new field. Creating catalogs, podcasts, and posts on social media can enhance awareness about the importance of this methodology to attract more participants to join this field of research.

- **Directions for data analysis developments.** Our analysis framework includes several advanced data analysis techniques to analyze spatio-temporal data. Yet, there are aspects overlooked in our proposed method that can be addressed in future works:

- *Tailoring spatio-temporal analysis to playground datasets.* More steps are required to make our proposed AI models functional at playgrounds. For example, access to representative data, i.e., data from different environmental designs from children with various age groups, capacities, and needs, is required to ensure that the designed model can perform reliably across various conditions and minimize the risk of harm to users or society. This could possibly be investigated in future works.
- *Developing automation process.* The design of the advanced AI models developed in this thesis highly depends on the characteristics of the given datasets. Adopting automated machine learning methods to design hyperparameter-free models can be investigated in the future.
- *Developing Data analysis techniques to tackle small sample sizes.* Advanced data analysis techniques might be able to address a small sample size problem. For example, developing GANs can be investigated in future research to create semi-synthesized samples similar to the original ones, thus creating a balanced and representative dataset from the research population.

7.4 Conclusion

This thesis aimed to develop a data analysis framework for capturing children's behavior in playgrounds. The proposed framework has addressed the three main characteristics of playgrounds, i.e., multiple environments, individual experiences, and spatio-temporal dynamics, by combining cutting-edge technological advancements in the light of the theory of affordances and a multidisciplinary collaboration involving psychology, architecture, and computer science experts. Specifically, our multimodal data-driven approach combined sensor data with other sources of information (e.g., observations, questionnaires, and self-reports) to ensure that children's interactions with the playground affordances have been accurately captured. Furthermore, by including individuals' experiences via self-report, our data analysis framework considered various effectivities in children. Lastly, the advanced data analytics and machine learning algorithms designed in this framework ensure that the spatio-temporal dynamics of group interactions have been correctly modeled.

Implementing this framework offers a valuable tool for academics, psychologists, sports clubs, policymakers, and designers. It enables them to analyze individual activities and identify individual—and group-level challenges and barriers in micro-communities. Moreover, this framework can validate a proposed intervention in micro-communities, e.g., adding a new design or setting a new policy.

CHAPTER 8

Appendix

Appendix A.

Correlations between the study variables (pooled and weighted results)

Variable	Spearman's rho in all children (in autistic/allistic children)					
	2.	3.	4.	5.	6.	7.
Age	-.275 (-.052/-.423)	.211 (.220/.129)	-.369 (-.177/-.445)	.248 (.282/.211)	-.239 (-.298/-.210)	.088 (.015/.113)
Loneliness	-	-.019 (.101/-.193)	.009 (-.344/.256)	-.128 (.268/-.437)	-.008 (-.035/.079)	.093 (.204/.020)
Reciprocated friendships	-	-	.079 (.152/.092)	.367 (.349/.346)	.072 (.154/.105)	.194 (.273/.133)
Classmate centrality	-	-	-	-.311 (-.367/-.236)	.276 (.159/.356)	-.133 (-.285/.015)
Time in social contact	-	-	-	-	.231 (.282/.206)	.495 (.692/.317)
N contact partners	-	-	-	-	-	.727 (.742/.735)
Playground centrality	-	-	-	-	-	-

Significant correlations ($p \leq .01$) are bolded.

Appendix B.

Results based on raw data (unweighted and before multiple imputation)

	Mean (SD)			Correlation with loneliness			
	Range	Autistic	Allistic	U	All		
Loneliness (total scorea; n = 85)	16-68	32.63 (10.02)	33.30 (11.92)	898.5	-	-	
Physical connectedness							
Total time in social contactb (n = 83)	0.03-1	.54 (.23)	.70 (.23)	522.0**	-.14	.11	
Number of contact partnersc (n = 83)	0.09-0.95	.60 (.17)	.53 (.17)	666.0	-.12	-	
Playground closeness centrality (n = 83)	0.02-0.11	.07 (.02)	.08 (.02)	741.0	-.02	-	
Emotional connectedness							
Reciprocated friendshipsd (n = 76)	0-1	.34 (.29)	.52 (.30)	459.5**	-.03	-	
Classmate closeness centrality (n = 75)	0.47-2	1.20 (.36)	.94 (.35)	433.5**	-.11	-	

Correlation coefficients for separate groups are reported only when Fisher's r-to-z transformation showed a significant difference in the strength of correlations between the group; otherwise, the correlation coefficients for the entire sample are reported.

[a] Highest possible total score is 80.

[b] Corrected by the total time when the child was detected.

[c] Corrected by n – 1, where n is the total number of children on the playground.

[d] Calculated as a degree by dividing the number of reciprocated nominations by the number of outgoing nominations.

$^{**}p \leq .01$.

$^{***}p \leq .001$.

Appendix C.

Results including only autistic children without comorbidity (n = 24) and allistic children without a diagnosis (n = 34) (weighted and pooled results)

		Mean (SD)			Correlation with loneliness		
	Range	**Autistic**	**Allistic**	**U**	**All**	**Autistic**	**Allistic**
Loneliness (total score[a])	16-68	33.32 (6.96)	35.99 (14.42)	687.50	-	-	-
Physical connectedness							
Total time in social contact[b]	0.03-0.99	.65 (.23)	.65 (.24)	600.00	-.07	.35	-.44**
Number of contact partners[c]	0.11-0.95	.53 (.14)	.54 (.15)	693.50	-.05	-	-
Playground closeness centrality	0.02-0.11	.08 (.02)	.08 (.02)	546.00	.15	-	-
Emotional connectedness							
Reciprocated friendships[d]	0-1	.46 (.26)	.49 (.28)	602.00	-.10	.28	-.36
Classmate closeness centrality	0.49-2	1.02 (.38)	.91 (.23)	533.50	-.20	-.43**	.23

Correlation coefficients for separate groups are reported only when Fisher's r-to-z transformation showed a significant difference in the strength of correlations between the group; otherwise, the correlation coefficients for the entire sample are reported.

[a] Highest possible total score is 80.

[b] Corrected by the total time when the child was detected.

[c] Corrected by n – 1, where n is the total number of children on the playground.

[d] Calculated as a degree by dividing the number of reciprocated nominations by the number of outgoing nominations.

$^{**}p \leq .01$.

$^{***}p \leq .001$.

Appendix D.

Mean and standard deviation of the study variables in autistic children with comorbidities (n = 23) and autistic children without comorbidities (n = 24) (weighted and pooled results).

	Autistic children with comorbidities Mean (SD)	**Autistic children without comorbidities** Mean (SD)	**U**	**p**
Loneliness	2c33.51 (10.25)	2c33.32 (6.96)	450.00	.932
Reciprocated friendships	2c.24 (.18)	2c.46 (.26)	238.50	.002
Classmate closeness centrality	2c1.12 (.24)	2c1.02 (.38)	415.00	.554
Total time in social contact	2c.56 (.20)	2c.65 (.23)	407.00	.478
Number of contact partners	2c.62 (.15)	2c.53 (.14)	295.50	.020
Playground closeness centrality	2c.08 (.02)	2c.08 (.02)	416.00	.564

References

[1] M. Nasri, Y.-T. Tsou, A. Koutamanis, M. Baratchi, S. Giest, D. Reidsma, and C. Rieffe, "A novel data-driven approach to examine children's movements and social behaviour in schoolyard environments," *Children*, vol. 9, no. 8, p. 1177, 2022.

[2] M. Nasri, M. Baratchi, Y.-T. Tsou, S. Giest, A. Koutamanis, and C. Rieffe, "A novel metric to measure spatio-temporal proximity: a case study analyzing children's social network in schoolyards," *Applied Network Science*, vol. 8, no. 1, p. 50, 2023.

[3] M. Nasri, Z. Fang, M. Baratchi, G. Englebienne, S. Wang, A. Koutamanis, and C. Rieffe, "A gnn-based architecture for group detection from spatio-temporal trajectory data," in *International Symposium on Intelligent Data Analysis*, pp. 327–339, Springer, 2023.

[4] M. Nasri, T. Maliappis, C. Rieffe, and M. Baratchi, "T-dante: Detecting group behaviour in spatio-temporal trajectories using context information," in *International Symposium on Intelligent Data Analysis*, pp. 28–39, Springer, 2024.

[5] Y.-T. Tsou, M. Nasri, B. Li, E. M. Blijd-Hoogewys, M. Baratchi, A. Koutamanis, and C. Rieffe, "Social connectedness and loneliness in school for autistic and allistic children," *Autism*, p. 13623613241259932, 2024.

[6] M. Nasri, S. Giest, Y.-T. Tsou, M. Baratchi, A. Koutamanis, and C. Rieffe, "Role of movement data in creating inclusive school settings for students," in *Data for Policy*, 2022.

[7] A. Eichengreen, M. van Rooijen, L.-M. van Klaveren, M. Nasri, Y.-T. Tsou, A. Koutamanis, M. Baratchi, and C. Rieffe, "The impact of loose-parts-play on schoolyard social participation of children with and without disabilities: A case study," *Child: care, health and development*, vol. 50, no. 1, p. e13144, 2024.

[8] A. Eichengreen, Y.-T. Tsou, M. Nasri, L.-M. van Klaveren, B. Li, A. Koutamanis, M. Baratchi, E. Blijd-Hoogewys, J. Kok, and C. Rieffe, "Social connectedness at the playground before and after covid-19 school closure," *Journal of applied developmental psychology*, vol. 87, p. 101562, 2023.

[9] R. Laanen, M. Nasri, R. van Dijk, M. Baratchi, A. Koutamanis, and C. Rieffe, "Automated classification of pre-defined movement patterns: A comparison between gnss and uwb technology," *arXiv preprint arXiv:2303.07107*, 2023.

[10] P. Diotaiuti, S. Mancone, and S. Corrado, "Using sports tracker: evidences on dependence, self-regulatory modes and resilience in a sample of competitive runners," *Psychology*, vol. 11, no. 1, pp. 54–70, 2020.

[11] N. De Raeve, A. Shahid, M. De Schepper, E. De Poorter, I. Moerman, J. Verhaevert, P. Van Torre, and H. Rogier, "Bluetooth-low-energy-based fall detection and warning system for elderly people in nursing homes," *Journal of Sensors*, vol. 2022, pp. 1–14, 2022.

[12] A. V. Creaser, S. A. Clemes, S. Costa, J. Hall, N. D. Ridgers, S. E. Barber, and D. D. Bingham, "The acceptability, feasibility, and effectiveness of wearable activity trackers for increasing physical activity in children and adolescents: a systematic review," *International journal of environmental research and public health*, vol. 18, no. 12, p. 6211, 2021.

[13] J. E. Druery, "Fostering sense of belonging: a multi-case study of black male retention initiatives.," *Electronic Theses and Dissertations.*, 2018.

[14] P. I. Karlsudd, "Inclusion through participation: Fostering pupils' feelings of belonging in swedish after-school care," *Education Sciences*, vol. 13, no. 4, p. 376, 2023.

[15] S. B. Sarason, "The psychological sense of community," *San Francisco*, 1974.

[16] J. E. van Dijk-Wesselius, A. E. Van den Berg, J. Maas, and D. Hovinga, "Green schoolyards as outdoor learning environments: Barriers and solutions as experienced by primary school teachers," *Frontiers in Psychology*, vol. 10, p. 484511, 2020.

[17] M. E. van den Berg, M. Winsall, S. M. Dyer, F. Breen, M. Gresham, and M. Crotty, "Understanding the barriers and enablers to using outdoor spaces in nursing homes: A systematic review," *The gerontologist*, vol. 60, no. 4, pp. e254–e269, 2020.

[18] L. Basterfield, L. Gardner, J. K. Reilly, M. S. Pearce, K. N. Parkinson, A. J. Adamson, J. J. Reilly, and S. A. Vella, "Can't play, won't play: longitudinal changes in perceived barriers to participation in sports clubs across the child–adolescent transition," *BMJ open sport & exercise medicine*, vol. 2, no. 1, p. e000079, 2016.

[19] R. J. Coplan, L. L. Ooi, A. Kirkpatrick, and K. H. Rubin, "Social and nonsocial play," in *Play from birth to twelve*, pp. 97–106, Routledge, 2015.

[20] F. Madondo and J. Tsikira, "Traditional children's games: their relevance on skills development among rural zimbabwean children age 3–8 years," *Journal of Research in Childhood Education*, vol. 36, no. 3, pp. 406–420, 2022.

[21] C. S. González, N. Gómez, V. Navarro, M. Cairós, C. Quirce, P. Toledo, and N. Marrero-Gordillo, "Learning healthy lifestyles through active videogames, motor games and the gamification of educational activities," *Computers in human behavior*, vol. 55, pp. 529–551, 2016.

[22] B. A. Cheatum and A. A. Hammond, *Physical activities for improving children's learning and behavior: A guide to sensory motor development*. Human Kinetics, 2000.

[23] A. J. Hanscom, *Balanced and barefoot: How unrestricted outdoor play makes for strong, confident, and capable children*. New Harbinger Publications, 2016.

[24] B. A. Corbett, C. W. Schupp, S. Levine, and S. Mendoza, "Comparing cortisol, stress, and sensory sensitivity in children with autism," *Autism research*, vol. 2, no. 1, pp. 39–49, 2009.

[25] A. E. Robertson and D. R. Simmons, "The relationship between sensory sensitivity and autistic traits in the general population," *Journal of Autism and Developmental disorders*, vol. 43, pp. 775–784, 2013.

[26] N. J. Minshew and J. A. Hobson, "Sensory sensitivities and performance on sensory perceptual tasks in high-functioning individuals with autism," *Journal of autism and developmental disorders*, vol. 38, pp. 1485–1498, 2008.

[27] J. G. Parker and S. R. Asher, "Friendship and friendship quality in middle childhood: Links with peer group acceptance and feelings of loneliness and social dissatisfaction.," *Developmental psychology*, vol. 29, no. 4, p. 611, 1993.

[28] J. Sparkes, "Schools, education and social exclusion," *LSE STICERD Research Paper No. CASE029*, 1999.

[29] N. Bauminger and C. Kasari, "Loneliness and friendship in high-functioning children with autism," *Child development*, vol. 71, no. 2, pp. 447–456, 2000.

[30] R. J. Coplan, L. L. Ooi, and L. Rose-Krasnor, "Naturalistic observations of schoolyard social participation: Marker variables for socio-emotional functioning in early adolescence," *The Journal of Early Adolescence*, vol. 35, no. 5-6, pp. 628–650, 2015.

[31] S. P. Hinshaw, C. Simmel, and T. L. Heller, "Multimethod assessment of covert antisocial behavior in children: Laboratory observations, adult ratings, and child self-report.," *Psychological Assessment*, vol. 7, no. 2, p. 209, 1995.

[32] T. Bouman, M. van der Meulen, F. A. Goossens, T. Olthof, M. M. Vermande, and E. A. Aleva, "Peer and self-reports of victimization and bullying: Their differential association with internalizing problems and social adjustment," *Journal of school psychology*, vol. 50, no. 6, pp. 759–774, 2012.

[33] D. Dessing, F. H. Pierik, R. P. Sterkenburg, P. van Dommelen, J. Maas, and S. I. de Vries, "Schoolyard physical activity of 6–11 year old children assessed by gps and accelerometry," *International Journal of Behavioral Nutrition and Physical Activity*, vol. 10, pp. 1–9, 2013.

[34] C. S. Pawlowski, H. B. Andersen, J. Troelsen, and J. Schipperijn, "Children's physical activity behavior during school recess: A pilot study using gps, accelerometer, participant observation, and go-along interview," *PloS one*, vol. 11, no. 2, p. e0148786, 2016.

[35] G. Veiga, W. De Leng, R. Cachucho, L. Ketelaar, J. N. Kok, A. Knobbe, C. Neto, and C. Rieffe, "Social competence at the playground: Preschoolers during recess," *Infant and Child Development*, vol. 26, no. 1, p. e1957, 2017.

[36] G. Veiga, L. Ketelaar, W. De Leng, R. Cachucho, J. N. Kok, A. Knobbe, C. Neto, and C. Rieffe, "Alone at the playground," *European Journal of Developmental Psychology*, vol. 14, no. 1, pp. 44–61, 2017.

[37] J. Amirian, B. Zhang, F. V. Castro, J. J. Baldelomar, J.-B. Hayet, and J. Pettré, "Opentraj: Assessing prediction complexity in human trajectories datasets," in *Proceedings of the asian conference on computer vision*, 2020.

[38] F. Aminpour, K. Bishop, and L. Corkery, "The hidden value of in-between spaces for children's self-directed play within outdoor school environments," *Landscape and urban planning*, vol. 194, p. 103683, 2020.

[39] J. L. Huberty, M. W. Beets, A. Beighle, and G. Welk, "Environmental modifications to increase physical activity during recess: preliminary findings from ready for recess," *Journal of Physical Activity and Health*, vol. 8, no. s2, pp. S249–S256, 2011.

[40] E. Baines and P. Blatchford, "Full report-school break and lunch times and young people's social lives: a follow-up national study," tech. rep., UCL Institute of Education, 2019.

[41] J. S. Owen-DeSchryver, E. G. Carr, S. I. Cale, and A. Blakeley-Smith, "Promoting social interactions between students with autism spectrum disorders and their peers in inclusive school settings," *Focus on Autism and other developmental disabilities*, vol. 23, no. 1, pp. 15–28, 2008.

[42] S. J. Slater, L. Nicholson, J. Chriqui, L. Turner, and F. Chaloupka, "The impact of state laws and district policies on physical education and recess practices in a nationally representative sample of us public elementary schools," *Archives of pediatrics & adolescent medicine*, vol. 166, no. 4, pp. 311–316, 2012.

[43] R. M. Stanley, K. Boshoff, and J. Dollman, "Voices in the playground: A qualitative exploration of the barriers and facilitators of lunchtime play," *Journal of science and medicine in sport*, vol. 15, no. 1, pp. 44–51, 2012.

[44] A. Jerebine, K. Fitton-Davies, N. Lander, E. L. Eyre, M. J. Duncan, and L. M. Barnett, ""all the fun stuff, the teachers say,'that's dangerous!'" hearing from children on safety and risk in active play in schools: a systematic," *Int J Behav Nutr Phys Act 19*, 2022.

[45] H. Heft, "Affordances of children's environments: A functional approach to environmental description," *Children's environments quarterly*, pp. 29–37, 1988.

[46] M. Kyttä, "Affordances of children's environments in the context of cities, small towns, suburbs and rural villages in finland and belarus," *Journal of environmental psychology*, vol. 22, no. 1-2, pp. 109–123, 2002.

[47] V. Sharma-Brymer, K. Davids, E. Brymer, and D. Bland, "Affordances in nature: Australian primary school children identify learning opportunities," *Curriculum Perspectives*, vol. 38, pp. 175–180, 2018.

[48] N. G. Cosco, R. C. Moore, and M. Z. Islam, "Behavior mapping: A method for linking preschool physical activity and outdoor design," *Medicine & Science in Sports & Exercise*, vol. 42, no. 3, pp. 513–519, 2010.

[49] W. R. Smith, R. Moore, N. Cosco, J. Wesoloski, T. Danninger, D. S. Ward, S. G. Trost, and N. Ries, "Increasing physical activity in childcare outdoor learning environments: The effect of setting adjacency relative to other built environment and social factors," *Environment and behavior*, vol. 48, no. 4, pp. 550–578, 2016.

[50] F. S. Flores, L. P. Rodrigues, F. Copetti, F. Lopes, and R. Cordovil, "Affordances for motor skill development in home, school, and sport environments: A narrative review," *Perceptual and motor skills*, vol. 126, no. 3, pp. 366–388, 2019.

[51] K. Bjørgen, "Physical activity in light of affordances in outdoor environments: qualitative observation studies of 3–5 years olds in kindergarten," *Springerplus*, vol. 5, no. 1, pp. 1–11, 2016.

[52] S. Everley and K. Everley, "Primary school children's experiences of physical activity: The place of social and cultural capital in participation and implications for schools," *Early Child Development and Care*, vol. 189, no. 12, pp. 2032–2042, 2019.

[53] A. Zaidman-Zait, P. Mirenda, E. Duku, T. Vaillancourt, I. M. Smith, P. Szatmari, S. Bryson, E. Fombonne, J. Volden, C. Waddell, *et al.*, "Impact of personal and social resources on parenting stress in mothers of children with autism spectrum disorder," *Autism*, vol. 21, no. 2, pp. 155–166, 2017.

[54] A. M. Bednarczyk, H. Alexander-Whiting, and G. A. Solit, "Guidelines for the adaptation of preschool environments to integrate deaf, hard of hearing, and hearing children," *Children's Environments*, pp. 6–15, 1994.

[55] S. D. Mayes, S. L. Calhoun, R. Aggarwal, C. Baker, S. Mathapati, S. Molitoris, and R. D. Mayes, "Unusual fears in children with autism," *Research in Autism Spectrum Disorders*, vol. 7, no. 1, pp. 151–158, 2013.

[56] A. Franklin, P. Sowden, R. Burley, L. Notman, and E. Alder, "Color perception in children with autism," *Journal of autism and developmental disorders*, vol. 38, pp. 1837–1847, 2008.

[57] Q. Wu, C. Yu, Y. Chen, J. Yao, X. Wu, X. Peng, and T. Han, "Squeeze the ball: Designing an interactive playground towards aiding social activities of children with low-function autism," in *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, pp. 1–14, 2020.

[58] M. Hudson, H. G. Burnett, and T. Jellema, "Anticipation of action intentions in autism spectrum disorder," *Journal of autism and developmental disorders*, vol. 42, pp. 1684–1693, 2012.

[59] D. W. Maynard, T. McDonald, and T. Stickle, "Parents as a team: Mother, father, a child with autism spectrum disorder, and a spinning toy," *Journal of autism and developmental disorders*, vol. 46, pp. 406–423, 2016.

[60] S. C. Wymer, J. Tarbox, G. A. Beavers, and C. A. Tullis, "Teaching children with autism to follow rules specifying a behavior and consequence," *The Analysis of Verbal Behavior*, vol. 32, pp. 265–274, 2016.

[61] P. Brambilla, A. Hardan, S. U. Di Nemi, J. Perez, J. C. Soares, and F. Barale, "Brain anatomy and development in autism: review of structural mri studies," *Brain research bulletin*, vol. 61, no. 6, pp. 557–569, 2003.

[62] S. M. Alessandri, "Attention, play, and social behavior in adhd preschoolers," *Journal of abnormal child psychology*, vol. 20, pp. 289–302, 1992.

[63] R. Cordier, A. Bundy, C. Hocking, and S. Einfeld, "A model for play-based intervention for children with adhd," *Australian Occupational Therapy Journal*, vol. 56, no. 5, pp. 332–340, 2009.

[64] E. E. Leipold and A. C. Bundy, "Playfulness in children with attention deficit hyperactivity disorder," *The Occupational Therapy Journal of Research*, vol. 20, no. 1, pp. 61–79, 2000.

[65] B. Da Silva, G. Veiga, and C. Rieffe, "A review on playground behaviors-the case of children with hearing loss," in *ESTUDOS EM DESENYOLYIMENTO MOTOR DA CRIANCA*, Edições Piaget, 2018.

[66] C. Rieffe, P. Oosterveld, M. M. Terwogt, S. Mootz, E. Van Leeuwen, and L. Stockmann, "Emotion regulation and internalizing symptoms in children with autism spectrum disorders," *Autism*, vol. 15, no. 6, pp. 655–670, 2011.

[67] B. F. Van Heijst and H. M. Geurts, "Quality of life in autism across the lifespan: A meta-analysis," *Autism*, vol. 19, no. 2, pp. 158–167, 2015.

[68] I. Fjørtoft, B. Kristoffersen, and J. Sageie, "Children in schoolyards: Tracking movement patterns and physical activity in schoolyards using global positioning system and heart rate monitoring," *Landscape and urban planning*, vol. 93, no. 3-4, pp. 210–217, 2009.

[69] A. P. Jones, E. G. Coombes, S. J. Griffin, and E. M. van Sluijs, "Environmental supportiveness for physical activity in english schoolchildren: a study using global positioning systems," *International Journal of Behavioral Nutrition and Physical Activity*, vol. 6, no. 1, pp. 1–8, 2009.

[70] R. Maddison and C. Ni Mhurchu, "Global positioning system: a new opportunity in physical activity measurement," *International journal of behavioral nutrition and physical activity*, vol. 6, no. 1, pp. 1–8, 2009.

[71] J. S. Duncan, H. M. Badland, and G. Schofield, "Combining gps with heart rate monitoring to measure physical activity in children: A feasibility study," *Journal of Science and Medicine in Sport*, vol. 12, no. 5, pp. 583–585, 2009.

[72] H. B. Andersen, L. B. Christiansen, C. S. Pawlowski, and J. Schipperijn, "What we build makes a difference–mapping activating schoolyard features after renewal using gis, gps and accelerometers," *Landscape and Urban Planning*, vol. 191, p. 103617, 2019.

[73] M. Kusenbach, "Street phenomenology: The go-along as ethnographic research tool," *Ethnography*, vol. 4, no. 3, pp. 455–485, 2003.

[74] M. G. Wing, A. Eklund, and L. D. Kellogg, "Consumer-grade global positioning system (gps) accuracy and reliability," *Journal of forestry*, vol. 103, no. 4, pp. 169–173, 2005.

[75] P. A. Zandbergen and S. J. Barbeau, "Positional accuracy of assisted gps data from high-sensitivity gps-enabled mobile phones," *The Journal of Navigation*, vol. 64, no. 3, pp. 381–399, 2011.

[76] G. Veiga and C. Rieffe, "Monitoring children's behaviors in their natural settings: Applying rfid sensors to study child development," *Sage Research Methods Cases*, vol. 10, p. 9781526446435, 2018.

[77] T. Elmer, K. Chaitanya, P. Purwar, and C. Stadtfeld, "The validity of rfid badges measuring face-to-face interactions," *Behavior research methods*, vol. 51, pp. 2120–2138, 2019.

[78] M. R. Puyau, A. L. Adolph, F. A. Vohra, and N. F. Butte, "Validation and calibration of physical activity monitors in children," *Obesity research*, vol. 10, no. 3, pp. 150–157, 2002.

[79] E. Delidou, O. Matsouka, and C. Nikolaidis, "Influence of school playground size and equipment on the physical activity of students during recess," *European Physical Education Review*, vol. 22, no. 2, pp. 215–224, 2016.

[80] M. Jansson, M. Abdulah, and A. Eriksson, "Secondary school students' perspectives and use of three school grounds of varying size, content and design," *Urban Forestry & Urban Greening*, vol. 30, pp. 115–123, 2018.

[81] L. Chawla, K. Keena, I. Pevec, and E. Stanley, "Green schoolyards as havens from stress and resources for resilience in childhood and adolescence," *Health & place*, vol. 28, pp. 1–13, 2014.

[82] E. Haug, T. Torsheim, J. F. Sallis, and O. Samdal, "The characteristics of the outdoor school environment associated with physical activity," *Health education research*, vol. 25, no. 2, pp. 248–256, 2010.

[83] M. Czalczynska-Podolska, "The impact of playground spatial features on children's play and activity forms: An evaluation of contemporary playgrounds' play and social value," *Journal of environmental psychology*, vol. 38, pp. 132–142, 2014.

[84] D. P. Loy, J. Dattilo, *et al.*, "Effects of different play structures on social interactions between a boy with asperger's syndrome and his peers," *Therapeutic Recreation Journal*, vol. 34, no. 3, pp. 190–210, 2000.

[85] L. E. Cohen, "Layers of discourse in preschool block play: An examination of children's social interactions," *International Journal of Early Childhood*, vol. 47, pp. 267–281, 2015.

[86] J. Van Dijk-Wesselius, J. Maas, D. Hovinga, M. Van Vugt, and A. Van den Berg, "The impact of greening schoolyards on the appreciation, and physical, cognitive and social-emotional well-being of schoolchildren: A prospective intervention study," *Landscape and urban planning*, vol. 180, pp. 15–26, 2018.

[87] C. Wang, B. Berry, and S. M. Swearer, "The critical role of school climate in effective bullying prevention," *Theory into practice*, vol. 52, no. 4, pp. 296–302, 2013.

[88] V. Djaja-Josko and J. Kolakowski, "Uwb positioning system for elderly persons monitoring," in *2015 23rd Telecommunications Forum Telfor (TELFOR)*, pp. 169–172, IEEE, 2015.

[89] A. Van den Bossche, R. Dalce, N. Gonzalez, and T. Val, "Locura: A new localisation and uwb-based ranging testbed for the internet of things," in *2018 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, pp. 1–6, IEEE, 2018.

[90] Y. Xu, Y. S. Shmaliy, C. K. Ahn, G. Tian, and X. Chen, "Robust and accurate uwb-based indoor robot localisation using integrated ekf/efir filtering," *IET radar, sonar & navigation*, vol. 12, no. 7, pp. 750–756, 2018.

[91] B. M. Heravi, J. L. Gibson, S. Hailes, and D. Skuse, "Playground social interaction analysis using bespoke wearable sensors for tracking and motion capture," in *Proceedings of the 5th International Conference on Movement and Computing*, pp. 1–8, 2018.

[92] D. Lusher, G. Robins, and P. Kremer, "The application of social network analysis to team sports," *Measurement in Physical Education and Exercise Science*, vol. 14, no. 4, pp. 211–224, 2010.

[93] X. Li, S. Geng, and S. Liu, "Social network analysis on tourists' perceived image of tropical forest park: Implications for niche tourism," *SAGE Open*, vol. 12, no. 1, p. 21582440211067243, 2022.

[94] M. Valeri and R. Baggio, "Social network analysis: Organizational implications in tourism management," *International Journal of Organizational Analysis*, vol. 29, no. 2, pp. 342–353, 2020.

[95] R. P. D. Redondo, C. Garcia-Rubio, A. F. Vilas, C. Campo, and A. Rodriguez-Carrion, "A hybrid analysis of lbsn data to early detect anomalies in crowd dynamics," *Future Generation Computer Systems*, vol. 109, pp. 83–94, 2020.

[96] J. A. Barnes and F. Harary, "Graph theory in network analysis," *Social Networks*, vol. 5, no. 2, pp. 235–244, 1983.

[97] A. Bavelas, "Communication patterns in task-oriented groups," *The Journal of The Acoustical Society of America*, vol. 22, no. 6, pp. 725–730, 1950.

[98] V. Kostakos, "Temporal graphs," *Physica A: Statistical Mechanics and its Applications*, vol. 388, no. 6, pp. 1007–1023, 2009.

[99] L. Franzini, M. N. Elliott, P. Cuccaro, M. Schuster, M. J. Gilliland, J. A. Grunbaum, F. Franklin, and S. R. Tortolero, "Influences of physical and social neighborhood environments on children's physical activity and obesity," *American Journal of Public Health*, vol. 99, no. 2, pp. 271–278, 2009.

[100] A. Montanari, C. Mascolo, K. Sailer, and S. Nawaz, "Detecting emerging activity-based working traits through wearable technology," *Proceedings of the International Conference (ACM) on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 1, no. 3, pp. 1–24, 2017.

[101] B. Chaix, "How daily environments and situations shape behaviors and health: momentary studies of mobile sensing and smartphone survey data," *Health & Place*, vol. 61, p. 102241, 2020.

[102] Z. Liu, X. Zhou, W. Shi, and A. Zhang, "Recommending attractive thematic regions by semantic community detection with multi-sourced vgi data," *International Journal of Geographical Information Science*, vol. 33, no. 8, pp. 1520–1544, 2019.

[103] E. Steiger, J. P. De Albuquerque, and A. Zipf, "An advanced systematic literature review on spatiotemporal analyses of twitter data," *Transactions in GIS*, vol. 19, no. 6, pp. 809–834, 2015.

[104] H. Advaith, S. Adarsh, G. Akshay, P. Sreeram, and G. Sajeev, "A proximity based community detection in temporal graphs," in *Proceedings of the International Conference on Electronics, Computing and Communication Technologies (CONECCT)*, pp. 1–6, IEEE, 2020.

[105] H. A. Rahmani, M. Aliannejadi, M. Baratchi, and F. Crestani, "A systematic analysis on the impact of contextual information on point-of-interest recommendation," *ACM Transactions on Information Systems (TOIS)*, vol. 40, no. 4, pp. 1–35, 2022.

[106] Y. Zheng, "Location-based social networks: Users," in *Computing with spatial trajectories*, pp. 243–276, New York: Springer, 2011.

[107] V. W. Zheng, Y. Zheng, X. Xie, and Q. Yang, "Collaborative location and activity recommendations with gps history data," in *Proceedings of the International Conference on World Wide Web*, pp. 1029–1038, 2010.

[108] Z. Ding, X. Li, C. Jiang, and M. Zhou, "Objectives and state-of-the-art of location-based social network recommender systems," *Acm Computing Surveys (Csur)*, vol. 51, no. 1, pp. 1–28, 2018.

[109] H. A. Rahmani, M. Aliannejadi, M. Baratchi, and F. Crestani, "Joint geographical and temporal modeling based on matrix factorization for point-of-interest recommendation," in *Proceedings of the European Conference on Information Retrieval*, pp. 205–219, Springer, 2020.

[110] H. Yoon, Y. Zheng, X. Xie, and W. Woo, "Smart itinerary recommendation based on user-generated gps trajectories," in *Proceedings of the International Conference on the Ubiquitous Intelligence and Computing (UIC)*, pp. 19–34, Springer, 2010.

[111] K. Vassakis, E. Petrakis, I. Kopanakis, J. Makridis, and G. Mastorakis, "Location-based social network data for tourism destinations," *Big Data and Innovation in Tourism, Travel, and Hospitality: Managerial Approaches, Techniques, and Applications*, pp. 105–114, 2019.

[112] I. P. Tussyadiah, "A concept of location-based social network marketing," *Journal of Travel & Tourism Marketing*, vol. 29, no. 3, pp. 205–220, 2012.

[113] L. Wang, X. Geng, X. Ma, F. Liu, and Q. Yang, "Crowd flow prediction by deep spatio-temporal transfer learning," *arXiv preprint arXiv:1802.00386*, 2018.

[114] I. Afyouni, A. Khan, and Z. A. Aghbari, "Deep-eware: spatio-temporal social event detection using a hybrid learning model," *Journal of Big Data*, vol. 9, no. 1, p. 86, 2022.

[115] L. Ruan, C. Li, Y. Zhang, and H. Wang, "Soft computing model based financial aware spatiotemporal social network analysis and visualization for smart cities," *Computers, Environment and Urban Systems*, vol. 77, p. 101268, 2019.

[116] C. Bracis, K. L. Bildstein, and T. Mueller, "Revisitation analysis uncovers spatio-temporal patterns in animal movement data," *Ecography*, vol. 41, no. 11, pp. 1801–1811, 2018.

[117] K.-Y. Yie, T.-W. Chien, Y.-T. Yeh, W. Chou, and S.-B. Su, "Using social network analysis to identify spatiotemporal spread patterns of covid-19 around the world: online dashboard development," *International Journal of Environmental Research and Public Health*, vol. 18, no. 5, p. 2461, 2021.

[118] W. Liu, B. Wang, Y. Yang, N. Mou, Y. Zheng, L. Zhang, and T. Yang, "Cluster analysis of microscopic spatio-temporal patterns of tourists' movement behaviors in mountainous scenic areas using open gps-trajectory data," *Tourism Management*, vol. 93, p. 104614, 2022.

[119] M. Baratchi, N. Meratnia, and P. J. Havinga, "Finding frequently visited paths: dealing with the uncertainty of spatio-temporal mobility data," in *2013 IEEE Eighth International Conference on Intelligent Sensors, Sensor Networks and Information Processing*, pp. 479–484, IEEE, 2013.

[120] D. B. West, *Introduction to graph theory*, vol. 2. London: Prentice hall Upper Saddle River, 2001.

[121] E. W. Dijksta, "A note on two problems in connexion with graphs," *Numerische mathematik*, vol. 1, no. 1, pp. 269–271, 1959.

[122] D. Müllner, "Modern hierarchical, agglomerative clustering algorithms," *arXiv preprint arXiv:1109.2378*, 2011.

[123] P. J. Rousseeuw, "Silhouettes: a graphical aid to the interpretation and validation of cluster analysis," *Journal of Computational and Applied Mathematics*, vol. 20, pp. 53–65, 1987.

[124] K. A. Clevenger, K. A. Pfeiffer, and A. L. Pearson, "Using linked accelerometer and gps data for characterizing children's schoolyard physical activity: An overview of hot spot analytic decisions with reporting guidelines," *Spatial and Spatio-temporal Epidemiology*, p. 100548, 2022.

[125] S. Schneider, A. Bolbos, J. Fessler, and C. Buck, "Deprivation amplification due to structural disadvantage? playgrounds as important physical activity resources for children and adolescents," *Public Health*, vol. 168, pp. 117–127, 2019.

[126] T. Caliński and J. Harabasz, "A dendrite method for cluster analysis," *Communications in Statistics-theory and Methods*, vol. 3, no. 1, pp. 1–27, 1974.

[127] D. L. Davies and D. W. Bouldin, "A cluster separation measure," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, no. 2, pp. 224–227, 1979.

[128] K. Allen, M. L. Kern, D. Vella-Brodrick, J. Hattie, and L. Waters, "What schools need to know about fostering school belonging: A meta-analysis," *Educational psychology review*, vol. 30, pp. 1–34, 2018.

[129] H. Craggs and C. Kelly, "Adolescents' experiences of school belonging: A qualitative meta-synthesis," *Journal of Youth Studies*, vol. 21, no. 10, pp. 1411–1425, 2018.

[130] S. Ascher and J. Paquette, "Loneliness and peer relations in childhood. currents directions in psychological science, 12, 75–78," 2003.

[131] J. T. Cacioppo, M. E. Hughes, L. J. Waite, L. C. Hawkley, and R. A. Thisted, "Loneliness as a specific risk factor for depressive symptoms: cross-sectional and longitudinal analyses.," *Psychology and aging*, vol. 21, no. 1, p. 140, 2006.

[132] Z. I. Santini, V. S. Pisinger, L. Nielsen, K. R. Madsen, M. K. Nelausen, A. Koyanagi, V. Koushede, S. Roffey, L. C. Thygesen, and C. Meilstrup, "Social disconnectedness, loneliness, and mental health among adolescents in danish high schools: A nationwide cross-sectional study," *Frontiers in behavioral neuroscience*, vol. 15, p. 632906, 2021.

[133] C. Chevallier, G. Kohls, V. Troiani, E. S. Brodkin, and R. T. Schultz, "The social motivation theory of autism," *Trends in cognitive sciences*, vol. 16, no. 4, pp. 231–239, 2012.

[134] G. Dawson, "Early behavioral intervention, brain plasticity, and the prevention of autism spectrum disorder," *Development and psychopathology*, vol. 20, no. 3, pp. 775–803, 2008.

[135] A. Klin, W. Jones, R. Schultz, and F. Volkmar, "The enactive mind, or from actions to cognition: lessons from autism," *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, vol. 358, no. 1430, pp. 345–360, 2003.

[136] A. Deckers, P. Muris, and J. Roelofs, "Being on your own or feeling lonely? loneliness and other social variables in youths with autism spectrum disorders," *Child Psychiatry & Human Development*, vol. 48, pp. 828–839, 2017.

[137] S. Bottini, "Social reward processing in individuals with autism spectrum disorder: A systematic review of the social motivation hypothesis," *Research in autism spectrum disorders*, vol. 45, pp. 9–26, 2018.

[138] V. K. Jaswal and N. Akhtar, "Being versus appearing socially uninterested: Challenging assumptions about social motivation in autism," *Behavioral and Brain Sciences*, vol. 42, p. e82, 2019.

[139] F. Rademaker, A. de Boer, E. Kupers, and A. Minnaert, "Applying the contact theory in inclusive education: A systematic review on the impact of contact and information on the social participation of students with disabilities," in *Frontiers in Education*, vol. 5, p. 602414, Frontiers Media SA, 2020.

[140] G. W. Allport, "The nature oj'prejudice," *Reading: Addison-Wesley*, vol. 129, 1954.

[141] A. Anderson, D. W. Moore, R. Godfrey, and C. M. Fletcher-Flinn, "Social skills assessment of children with autism in free-play situations," *Autism*, vol. 8, no. 4, pp. 369–385, 2004.

[142] Y.-C. Chang, C.-H. Chen, P.-C. Huang, and L.-Y. Lin, "Understanding the characteristics of friendship quality, activity participation, and emotional well-being in taiwanese adolescents with autism spectrum disorder," *Scandinavian journal of occupational therapy*, vol. 26, no. 6, pp. 452–462, 2019.

[143] M. Falkmer, K. Oehlers, M. Granlund, and T. Falkmer, "Can you see it too? observed and self-rated participation in mainstream schools in students with and without autism spectrum disorders," *Developmental neurorehabilitation*, vol. 18, no. 6, pp. 365–374, 2015.

[144] J. Locke, W. Shih, M. Kretzmann, and C. Kasari, "Examining playground engagement between elementary school children with and without autism spectrum disorder," *Autism*, vol. 20, no. 6, pp. 653–662, 2016.

[145] V. Phares, "Father absence, mother love, and other family issues that need to be questioned: Comment on silverstein (1993).," *Journal of Family Psychology*, 1993.

[146] T. Bowles and J. Scull, "The centrality of connectedness: A conceptual synthesis of attending, belonging, engaging and flowing," *Journal of Psychologists and Counsellors in Schools*, vol. 29, no. 1, pp. 3–21, 2019.

[147] K. Bottema-Beutel, C. Malloy, J. Cuda, S. Y. Kim, and J. P. MacEvoy, "Friendship expectations may be similar for mental age-matched children with autism spectrum disorder and typically developing children," *Journal of autism and developmental disorders*, vol. 49, pp. 4346–4354, 2019.

[148] B. Chamberlain, C. Kasari, and E. Rotheram-Fuller, "Involvement or isolation? the social networks of children with autism in regular classrooms," *Journal of autism and developmental disorders*, vol. 37, pp. 230–242, 2007.

[149] C. Kasari, J. Locke, A. Gulsrud, and E. Rotheram-Fuller, "Social networks and friendships at school: Comparing children with and without asd," *Journal of autism and developmental disorders*, vol. 41, pp. 533–544, 2011.

[150] J. Locke, C. Kasari, E. Rotheram-Fuller, M. Kretzmann, and J. Jacobs, "Social network changes over the school year among elementary school-aged children with and without an autism spectrum disorder," *School Mental Health*, vol. 5, pp. 38–47, 2013.

[151] M. Dean, C. Kasari, W. Shih, F. Frankel, R. Whitney, R. Landa, C. Lord, F. Orlich, B. King, and R. Harwood, "The peer relationships of girls with asd at school: comparison to boys and girls with and without asd," *Journal of Child Psychology and Psychiatry*, vol. 55, no. 11, pp. 1218–1225, 2014.

[152] L. Calder, V. Hill, and E. Pellicano, "'sometimes i want to play by myself': Understanding what friendship means to children with autism in mainstream primary schools," *Autism*, vol. 17, no. 3, pp. 296–316, 2013.

[153] S. O'Hagan and J. Hebron, "Perceptions of friendship among adolescents with autism spectrum conditions in a mainstream high school resource provision," *European Journal of Special Needs Education*, vol. 32, no. 3, pp. 314–328, 2017.

[154] F. Sedgewick, V. Hill, R. Yates, L. Pickering, and E. Pellicano, "Gender differences in the social motivation and friendship experiences of autistic and non-autistic adolescents," *Journal of autism and developmental disorders*, vol. 46, pp. 1297–1306, 2016.

[155] R. S. V. Foggo and A. A. Webster, "Understanding the social experiences of adolescent females on the autism spectrum," *Research in Autism Spectrum Disorders*, vol. 35, pp. 74–85, 2017.

[156] L. Cresswell, R. Hinch, and E. Cage, "The experiences of peer relationships amongst autistic adolescents: A systematic review of the qualitative evidence," *Research in Autism Spectrum Disorders*, vol. 61, pp. 45–60, 2019.

[157] N. Bauminger, C. Shulman, and G. Agam, "Peer interaction and loneliness in high-functioning children with autism," *Journal of autism and developmental disorders*, vol. 33, pp. 489–507, 2003.

[158] J. Locke, E. H. Ishijima, C. Kasari, and N. London, "Loneliness, friendship quality and the social networks of adolescents with high-functioning autism in an inclusive school setting," *Journal of research in special educational needs*, vol. 10, no. 2, pp. 74–81, 2010.

[159] A. J. Whitehouse, K. Durkin, E. Jaquet, and K. Ziatas, "Friendship, loneliness and depression in adolescents with asperger's syndrome," *Journal of adolescence*, vol. 32, no. 2, pp. 309–322, 2009.

[160] T. J. Berndt, "The features and effects of friendship in early adolescence," *Child development*, pp. 1447–1460, 1982.

[161] C. Cattuto, W. Van den Broeck, A. Barrat, V. Colizza, J.-F. Pinton, and A. Vespignani, "Dynamics of person-to-person interactions from distributed rfid sensor networks," *PloS one*, vol. 5, no. 7, p. e11596, 2010.

[162] L. C. Freeman *et al.*, "Centrality in social networks: Conceptual clarification," *Social network: critical concepts in sociology. Londres: Routledge*, vol. 1, pp. 238–263, 2002.

[163] J. Zhang and Y. Luo, "Degree centrality, betweenness centrality, and closeness centrality in social network," in *2017 2nd international conference on modelling, simulation and applied mathematics (MSAM2017)*, pp. 300–303, Atlantis press, 2017.

[164] S. J. Pijl, P. Frostad, and A. Flem, "The social position of pupils with special needs in regular schools," *Scandinavian Journal of Educational Research*, vol. 52, no. 4, pp. 387–405, 2008.

[165] S. R. Asher and V. A. Wheeler, "Children's loneliness: a comparison of rejected and neglected peer status.," *Journal of consulting and clinical psychology*, vol. 53, no. 4, p. 500, 1985.

[166] M. Maes, W. Van den Noortgate, J. Vanhalst, W. Beyers, and L. Goossens, "The children's loneliness scale: Factor structure and construct validity in belgian children," *Assessment*, vol. 24, no. 2, pp. 244–251, 2017.

[167] G. Csardi and T. Nepusz, "The igraph software," *Complex syst*, vol. 1695, pp. 1–9, 2006.

[168] R. C. Team *et al.*, "R: A language and environment for statistical computing," *Foundation for Statistical Computing, Vienna, Austria*, 2013.

[169] G. Van Rossum and F. L. Drake Jr, "Python tutorial," 1995.

[170] M. J. Azur, E. A. Stuart, C. Frangakis, and P. J. Leaf, "Multiple imputation by chained equations: what is it and how does it work?," *International journal of methods in psychiatric research*, vol. 20, no. 1, pp. 40–49, 2011.

[171] J. L. Schafer and J. W. Graham, "Missing data: our view of the state of the art.," *Psychological methods*, vol. 7, no. 2, p. 147, 2002.

[172] J. R. Van Ginkel, M. Linting, R. C. Rippe, and A. van der Voort, "Rebutting existing misconceptions about multiple imputation as a method for handling missing data," *Journal of personality assessment*, vol. 102, no. 3, pp. 297–308, 2020.

[173] P. C. Austin and E. A. Stuart, "Moving towards best practice when using inverse probability of treatment weighting (iptw) using the propensity score to estimate causal treatment effects in observational studies," *Statistics in medicine*, vol. 34, no. 28, pp. 3661–3679, 2015.

[174] N. C. Chesnaye, V. S. Stel, G. Tripepi, F. W. Dekker, E. L. Fu, C. Zoccali, and K. J. Jager, "An introduction to inverse probability of treatment weighting in observational research," *Clinical Kidney Journal*, vol. 15, no. 1, pp. 14–20, 2022.

[175] C. Ruth, M. Brownell, J. Isbister, L. MacWilliam, H. Gammon, D. Singal, and E. Boreskewich, *Long-term outcomes of manitoba's insight mentoring program: a comparative statistical analysis*. Manitoba Centre for Health Policy Winnipeg, MB, 2015.

[176] P. Blatchford, P. Bassett, and P. Brown, "Examining the effect of class size on classroom engagement and teacher–pupil interaction: Differences in relation to pupil prior attainment and primary vs. secondary schools," *Learning and instruction*, vol. 21, no. 6, pp. 715–730, 2011.

[177] F. de Swart, W. J. Burk, W. B. Nelen, and R. H. Scholte, "Peer preference, perceived popularity, and the teacher–child relationship in special education," *Remedial and Special Education*, vol. 42, no. 2, pp. 67–77, 2021.

[178] K. L. Lane, J. H. Wehby, M. A. Little, and C. Cooley, "Academic, social, and behavioral profiles of students with emotional and behavioral disorders educated in self-contained classrooms and self-contained schools: Part i—are they more alike than different?," *Behavioral Disorders*, vol. 30, no. 4, pp. 349–361, 2005.

[179] L. Sokal and U. Sharma, "Canadian in-service teachers' concerns, efficacy, and attitudes about inclusive teaching," *Exceptionality Education International*, vol. 23, no. 1, 2014.

[180] L. Florian, "Reimagining special education: Why new approaches are needed," *The SAGE handbook of special education*, vol. 1, pp. 9–22, 2014.

[181] C. Rieffe and A. Koutamanis, "Een autisme-vriendelijke school: Architectuur als eerste stap," *Wetenschappelijk Tijdschrift Autisme*, vol. 22, no. 4, pp. 29–39, 2023.

[182] M. Mostafa, "Architecture for autism: Autism aspectss™ in school design," *International Journal of Architectural Research: ArchNet-IJAR*, vol. 8, no. 1, pp. 143–158, 2014.

[183] K. Harris, P. Rosinski, J. Wood-Nartker, and R. Hill Renirie, "Developing inclusive playgrounds that welcome all children—including those with autism," *Review Journal of Autism and Developmental Disorders*, pp. 1–9, 2022.

[184] N. Yuill, S. Strieth, C. Roake, R. Aspden, and B. Todd, "Brief report: Designing a playground for children with autistic spectrum disorders—-effects on playful peer interactions," *Journal of autism and developmental disorders*, vol. 37, pp. 1192–1196, 2007.

[185] M. H. Black, R. Kuzminski, J. Wang, J. Ang, C. Lee, S. Hafidzuddin, and S. McGarry, "Experiences of friendships for individuals on the autism spectrum: A scoping review," *Review Journal of Autism and Developmental Disorders*, pp. 1–26, 2022.

[186] B. A. Corbett, D. M. Swain, C. Newsom, L. Wang, Y. Song, and D. Edgerton, "Biobehavioral profiles of arousal and social motivation in autism spectrum disorders," *Journal of child psychology and psychiatry*, vol. 55, no. 8, pp. 924–934, 2014.

[187] M. Dean, R. Harwood, and C. Kasari, "The art of camouflage: Gender differences in the social behaviors of girls and boys with autism spectrum disorder," *Autism*, vol. 21, no. 6, pp. 678–689, 2017.

[188] A. Mannion and G. Leader, "Comorbidity in autism spectrum disorder: A literature review," *Research in Autism Spectrum Disorders*, vol. 7, no. 12, pp. 1595–1616, 2013.

[189] E. Holmes and T. Willoughby, "Play behaviour of children with autism spectrum disorders," *Journal of Intellectual and Developmental Disability*, vol. 30, no. 3, pp. 156–164, 2005.

[190] K. S. Gunn, D. Trembath, and K. Hudry, "An examination of interactions among children with autism and their typically developing peers," *Developmental Neurorehabilitation*, vol. 17, no. 5, pp. 327–338, 2014.

[191] D. UNICEF *et al.*, "Seen, counted, included: Using data to shed light on the well-being of children with disabilities," *New York: United Nations Children's Fund*, 2021.

[192] A. Ermagun and D. Levinson, "Spatiotemporal traffic forecasting: review and proposed directions," *Transport Reviews*, vol. 38, no. 6, pp. 786–814, 2018.

[193] Y. Chang, J. Qi, Y. Liang, and E. Tanin, "Contrastive trajectory similarity learning with dual-feature attention," in *2023 IEEE 39th International Conference on Data Engineering (ICDE)*, pp. 2933–2945, IEEE, 2023.

[194] D. Yao, C. Zhang, Z. Zhu, J. Huang, and J. Bi, "Trajectory clustering via deep representation learning," in *2017 international joint conference on neural networks (IJCNN)*, pp. 3880–3887, IEEE, 2017.

[195] X. Li, K. Zhao, G. Cong, C. S. Jensen, and W. Wei, "Deep representation learning for trajectory similarity computation," in *2018 IEEE 34th international conference on data engineering (ICDE)*, pp. 617–628, IEEE, 2018.

[196] L. Liu, J. Zhen, G. Li, G. Zhan, Z. He, B. Du, and L. Lin, "Dynamic spatial-temporal representation learning for traffic flow prediction," *IEEE Transactions on Intelligent Transportation Systems*, vol. 22, no. 11, pp. 7169–7183, 2020.

[197] S. Dabiri, C.-T. Lu, K. Heaslip, and C. K. Reddy, "Semi-supervised deep learning approach for transportation mode identification using gps trajectory data," *IEEE Transactions on Knowledge and Data Engineering*, vol. 32, no. 5, pp. 1010–1023, 2019.

[198] T.-Y. Fu and W.-C. Lee, "Trembr: Exploring road networks for trajectory representation learning," *ACM Transactions on Intelligent Systems and Technology (TIST)*, 2020.

[199] J. Jiang, D. Pan, H. Ren, X. Jiang, C. Li, and J. Wang, "Self-supervised trajectory representation learning with temporal regularities and travel semantics," in *2023 IEEE 39th international conference on data engineering (ICDE)*, IEEE, 2023.

[200] X. Liu, Y. Liang, C. Huang, Y. Zheng, B. Hooi, and R. Zimmermann, "When do contrastive learning signals help spatio-temporal graph forecasting?," in *Proceedings of the 30th International Conference on Advances in Geographic Information Systems*, pp. 1–12, 2022.

[201] Z. Fan, X. Song, Q. Chen, R. Jiang, R. Shibasaki, and K. Tsubouchi, "Trajectory fingerprint: one-shot human trajectory identification using siamese network," *CCF Transactions on Pervasive Computing and Interaction*, vol. 2, no. 2, pp. 113–125, 2020.

[202] F. Schroff, D. Kalenichenko, and J. Philbin, "Facenet: A unified embedding for face recognition and clustering," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 815–823, 2015.

[203] A. v. d. Oord, Y. Li, and O. Vinyals, "Representation learning with contrastive predictive coding," *arXiv preprint arXiv:1807.03748*, 2018.

[204] J. Bromley, J. W. Bentz, L. Bottou, I. Guyon, Y. LeCun, C. Moore, E. Säckinger, and R. Shah, "Signature verification using a "siamese" time delay neural network," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 7, no. 04, 1993.

[205] A. Yan, S. Cheng, W.-C. Kang, M. Wan, and J. McAuley, "Cosrec: 2d convolutional neural networks for sequential recommendation," in *Proceedings of the 28th ACM international conference on information and knowledge management*, 2019.

[206] J.-F. Chen, W.-L. Chen, C.-P. Huang, S.-H. Huang, and A.-P. Chen, "Financial time-series data analysis using deep convolutional neural networks," in *2016 7th International conference on cloud computing and big data (CCBD)*, pp. 87–92, IEEE, 2016.

[207] Y. Sun, C. Cheng, Y. Zhang, C. Zhang, L. Zheng, Z. Wang, and Y. Wei, "Circle loss: A unified perspective of pair similarity optimization," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020.

[208] F. Wilcoxon, "Individual comparisons by ranking methods. biometrics bulletin. 1945; 1 (6): 80–83," *International Biometric Society, Wiley, US*, 2012.

[209] S. Pellegrini, A. Ess, K. Schindler, and L. Van Gool, "You'll never walk alone: Modeling social behavior for multi-target tracking," in *2009 IEEE 12th International Conference on Computer Vision*, pp. 261–268, IEEE, 2009.

[210] A. Lerner, Y. Chrysanthou, and D. Lischinski, "Crowds by example," in *Computer graphics forum*, vol. 26, pp. 655–664, Wiley Online Library, 2007.

[211] A. Strehl and J. Ghosh, "Cluster ensembles—a knowledge reuse framework for combining multiple partitions," *Journal of machine learning research*, vol. 3, no. Dec, 2002.

[212] E. B. Fowlkes and C. L. Mallows, "A method for comparing two hierarchical clusterings," *Journal of the American statistical association*, vol. 78, no. 383, pp. 553–569, 1983.

[213] J. MacQueen *et al.*, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, vol. 1, pp. 281–297, Oakland, CA, USA, 1967.

[214] P. Khosla, P. Teterwak, C. Wang, A. Sarna, Y. Tian, P. Isola, A. Maschinot, C. Liu, and D. Krishnan, "Supervised contrastive learning," *Advances in neural information processing systems*, vol. 33, pp. 18661–18673, 2020.

[215] J.-Y. Franceschi, A. Dieuleveut, and M. Jaggi, "Unsupervised scalable representation learning for multivariate time series," *Advances in neural information processing systems*, vol. 32, 2019.

[216] J. Juvonen, L. M. Lessard, R. Rastogi, H. L. Schacter, and D. S. Smith, "Promoting social inclusion in educational settings: Challenges and opportunities," *Educational Psychologist*, vol. 54, no. 4, pp. 250–270, 2019.

[217] M. F. Hazlehurst, K. L. Wolf, C. Simmons, C. Nieto, M. K. Steiner, K. A. Garrett, A. V. Faino, M. Ubalde López, M. López-Toribio, and P. S. Tandon, "Physical activity and social interaction assessments in schoolyard settings using the system for observing outdoor play environments in neighborhood schools (soopen)," *International Journal of Behavioral Nutrition and Physical Activity*, vol. 20, no. 1, p. 94, 2023.

[218] J. Stehlé, N. Voirin, A. Barrat, C. Cattuto, L. Isella, J.-F. Pinton, M. Quaggiotto, W. Van den Broeck, C. Régis, B. Lina, *et al.*, "High-resolution measurements of face-to-face contact patterns in a primary school," *PloS one*, vol. 6, no. 8, p. e23176, 2011.

[219] M. Hoffman, P. Block, T. Elmer, and C. Stadtfeld, "A model for the dynamics of face-to-face interactions in social groups," *Network Science*, vol. 8, no. S1, pp. S4–S25, 2020.

[220] X. Que, F. Checconi, F. Petrini, and J. A. Gunnels, "Scalable community detection with the louvain algorithm," in *2015 IEEE international parallel and distributed processing symposium*, pp. 28–37, IEEE, 2015.

[221] J. Yang, J. McAuley, and J. Leskovec, "Community detection in networks with node attributes," in *2013 IEEE 13th international conference on data mining*, pp. 1151–1156, IEEE, 2013.

[222] V. D. Blondel, J.-L. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2008, no. 10, p. P10008, 2008.

[223] F. Solera, S. Calderara, and R. Cucchiara, "Socially constrained structural learning for groups detection in crowd," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 5, pp. 995–1008, 2015.

[224] K. Yamaguchi, A. C. Berg, L. E. Ortiz, and T. L. Berg, "Who are you with and where are you going?," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1345–1352, IEEE, 2011.

[225] M. Swofford, J. Peruzzi, N. Tsoi, S. Thompson, R. Martín-Martín, S. Savarese, and M. Vázquez, "Improving social awareness through dante: Deep affinity network for clustering conversational interactants," *Proc. ACM Hum.-Comput. Interact.*, vol. 4, may 2020.

[226] T. Kipf, E. Fetaya, K.-C. Wang, M. Welling, and R. Zemel, "Neural relational inference for interacting systems," in *Proceedings of the International Conference on Machine Learning*, pp. 2688–2697, PMLR, 2018.

[227] T. Fernando, S. Denman, S. Sridharan, and C. Fookes, "Gd-gan: Generative adversarial networks for trajectory prediction and group detection in crowds," in *Computer Vision–ACCV 2018: 14th Asian Conference on Computer Vision, Perth, Australia, December 2–6, 2018, Revised Selected Papers, Part I 14*, pp. 314–330, Springer, 2019.

[228] J. Li, R. Han, H. Yan, Z. Qian, W. Feng, and S. Wang, "Self-supervised social relation representation for human group detection," in *European Conference on Computer Vision*, pp. 142–159, Springer, 2022.

[229] M. Nasri, Z. Fang, M. Baratchi, G. Englebienne, S. Wang, A. Koutamanis, and C. Rieffe, "A gnn-based architecture for group detection from spatio-temporal trajectory data," in *Advances in Intelligent Data Analysis XXI* (B. Crémilleux, S. Hess, and S. Nijssen, eds.), (Cham), pp. 327–339, Springer Nature Switzerland, 2023.

[230] S. Thompson, A. Gupta, A. W. Gupta, A. Chen, and M. Vázquez, "Conversational group detection with graph neural networks," in *Proceedings of the International Conference on Multimodal Interaction*, pp. 248–252, 2021.

[231] A. v. d. Oord, S. Dieleman, H. Zen, K. Simonyan, O. Vinyals, A. Graves, N. Kalchbrenner, A. Senior, and K. Kavukcuoglu, "Wavenet: A generative model for raw audio," *arXiv preprint arXiv:1609.03499*, 2016.

[232] X. Lu, X. Li, C. Hu, J. Deng, W. Sheng, and L. Zhu, "A two-branch deep learning with spatial and pose constraints for social group detection," *Engineering Applications of Artificial Intelligence*, vol. 124, p. 106583, 2023.

[233] H. Hung and B. Kröse, "Detecting f-formations as dominant sets," in *Proceedings of the 13th International Conference on Multimodal Interfaces*, ICMI '11, (New York, NY, USA), p. 231–238, Association for Computing Machinery, 2011.

[234] J. Amirian, B. Zhang, F. V. Castro, J. J. Baldelomar, J.-B. Hayet, and J. Pettre, "Opentraj: Assessing prediction complexity in human trajectories datasets," in *Asian Conference on Computer Vision (ACCV)*, Springer, 2020.

[235] F. Solera, S. Calderara, and R. Cucchiara, "Socially constrained structural learning for groups detection in crowd," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 38, no. 5, pp. 995–1008, 2016.

[236] M. Vilain, J. D. Burger, J. Aberdeen, D. Connolly, and L. Hirschman, "A model-theoretic coreference scoring scheme," in *Sixth Message Understanding Conference (MUC-6): Proceedings of a Conference Held in Columbia, Maryland, November 6-8, 1995*, 1995.

[237] T. Fernando, S. Denman, S. Sridharan, and C. Fookes, "Gd-gan: Generative adversarial networks for trajectory prediction and group detection in crowds," in *Computer Vision – ACCV 2018* (C. V. Jawahar, H. Li, G. Mori, and K. Schindler, eds.), (Cham), pp. 314–330, Springer International Publishing, 2019.

[238] I. Bae, J.-H. Park, and H.-G. Jeon, "Learning pedestrian group representations for multimodal trajectory prediction," in *European Conference on Computer Vision*, pp. 270–289, Springer, 2022.

[239] R. F. Woolson, “Wilcoxon signed-rank test,” *Wiley encyclopedia of clinical trials*, pp. 1–3, 2007.

[240] B. R. Fajen, M. A. Riley, and M. T. Turvey, “Information, affordances, and the control of action in sport,” *international Journal of sport psychology*, vol. 40, no. 1, pp. 79–107, 2008.

Acknowledgments

I want to express my deepest gratitude to all those who have contributed to completing this dissertation.

First and foremost, I want to express my appreciation to all the children, parents, teachers, and schools who participated in and contributed their time to this research. Without their cooperation, this work would not have been possible.

I am thankful to my supervisors, Prof.dr. Carolien Rieffe, Dr. Mitra Baratchi, and Dr. Alexander Koutamanis, for giving me the opportunity to do this multidisciplinary research. Carolien, thank you for all the opportunities you have given me and your invaluable insights into various topics and many life lessons. Mitra, your knowledge and all the excellent discussions on our analyses always provide me with refreshing thoughts. Thanks for your unwavering guidance and patience throughout all my research papers. Alexander, I have learned a lot from our Monday afternoon meetings, from the world of affordances to all the life lessons. Thanks for all your guidance throughout my PhD journey. I am also thankful to Prof.dr. Sarah Giest for her contribution and input to my research.

I am grateful to my friends and colleagues at the Focus on Emotions lab for all their support, lunchtime, and lab meeting discussions. Yung-Ting, Boya, Brenda, Jiyain, Kexin, Claudia, Meng Qi, Zijian, Shannon, Lisa, Charlotte, and Lia, I have learned a lot from each one of you. Yung-Ting, talking to you and discussing my struggles and uncertainties was always helpful. Thanks for always being there. Thanks to Nina, Iris, and Chang for being the best office mates and making all the pleasant memories on Tuesdays. My paranymphs, Kexin and Claudia, thank you for helping me organize the defense ceremony.

Thanks to all my friends in the Netherlands, Haleh Hayati, Majid Ghorbani, Sahar Safarloo, Abbas Jamali, Sahar Hosseini, Amirreza Yousefzadeh, Zahra Forouzi, Amirabbas Kashi, little Ario, Farimah Ghassemi, Mehdi Karamzadeh, Faezeh Kardan, Babak Mohammadi, Parisa Ghanooni, Hadi Balaghi, Sanaz Lari, Parastoo, Morteza Moradi, Parvin Kolahkaj, and Ahmad Hadi, who made me feel at home with all Shelem nights, volleyball evenings and campings.

Thanks to all my old friends, Zahra Alipour, Yalda Foroutan, Anahita Baninajar, Setareh Rahimi, and Mahsa Khosravi, for your support, no matter the distance between us now.

Thanks to my mother-in-law and father-in-law, Fati and Reza, sisters-in-law Elnaz and Anahita, brother-in-law Arash, and sweet nephew-in-law Adrian for all your encouragement throughout this journey.

My sincere gratitude goes to my family for their belief in me. Thanks, Baba and Maman, for always having my back in all my decisions. Your love and support have been my pillars of strength. My beloved sister, Mahsa, and my brother-in-law, Hanif, thank you for always supporting me at every stage, and my sweet little niece, Waniya, for bringing joy and happiness into our lives.

Ashkan, my husband and multi-aspect life partner, it would have been impossible without you. Thanks for all your love, support, and guidance.

Thanks to everyone who stood part of this academic journey and supported me in reaching this significant milestone.

Maedeh Nasri,
Leiden, June 2024.

About the Author

Maedeh Nasri was born in Ramsar, Iran, in 1994. She finished high school in mathematics and physics in Farzanegan, National Organization for Development of Exceptional Talents. In 2012, she started her B.Sc. in the Electrical Engineering department at the Polytechnic of Tehran, Iran. After completing the bachelor program in 2016, she was admitted as an exceptionally talented student to the M.Sc. study in Electrical Engineering—control systems with exemption from the M.Sc. entrance exam at the same university. She finished her M.Sc. in Electrical Engineering under the supervision of Prof. H.A. Talebi and Dr. M.A. Khosravi.

In 2018, she moved to the Netherlands to start her career as a test engineer in Atlas Copco B.V., Oosterhout, during which she earned the Outstanding Individual Contribution Award from Atlas Copco.

While looking for a unique challenge, in 2020, she started her Ph.D. program at Leiden University under the supervision of Prof. Carolien Rieffe and Dr. Mitra Baratchi at Leiden University and the supervision of Dr. Alexander Koutamanis from TU Delft. Her PhD project, "Breaking the Cycle", was jointly funded by NWO and the center of BOLD cities, focused on adopting modern sensing technologies and advanced data analytics techniques to identify spatio-temporal patterns. During her Ph.D., she published several conference and journal papers and earned the Wild Card Award for "Interdisciplinary Research" from Leiden University in 2022.