A Framework for Improving Social Inclusion using Network Analysis and IoT-based Contact Tracing

Saverio Giallorenzo[1,2], Emanuele Sinagra[1], Angelo Trotta[1], Loris Vergolini[3,4], Andrea Zanellati[1]
[1]Dipartimento di Informatica - Scienza e Ingegneria, Università di Bologna, Bologna, Italy
[2]OLAS Team, INRIA, Sophia Antipolis, France
[3]Dipartimento di Scienze Politiche e Sociali, Università di Bologna, Bologna, Italy
[4]FBK-IRVAPP, Italy
{saverio.giallorenzo, angelo.trotta5, loris.vergolini, andrea.zanellati2}@unibo.it,
emanuelesinagra@gmail.com

Abstract—Social isolation poses critical challenges, with profound implications for performance, mental health, and general well-being. We present a framework to quantitatively measure social inclusion and suggest actions to increase the integration of isolated individuals. Specifically, we address the problem of detection and mitigation of social isolation in educational contexts – a pressing concern where the fundamental role of peer relationships contributes to determining student health and academic success. To tackle this challenge, we employ affordable, off-the-shelf IoT devices for reliably detecting face-to-face social interactions. On the detected network, we employ network analysis techniques, particularly PageRank and Betweenness centrality, to propose a novel algorithm that infers social inclusion levels and recommends sustainable interventions aligned with natural interaction patterns, thus ensuring sustainable integration. Our approach addresses the growing need for evidence-based, scalable solutions implementable across diverse educational environments without substantial infrastructure investment. We validate the framework through a real-world case study in a primary school, demonstrating the effectiveness of our methods in identifying socially isolated students and providing actionable insights for their social integration, offering an affordable tool for addressing one of education's most persistent challenges.

Index Terms—Social Inclusion, IoT, Social Network Analysis, Bluetooth Low Energy, Empirical Study

1 Introduction

Diversity, Equity, and Inclusion (DEI) are gaining adoption in public and private environments, like workplaces and schools. Diversity is now an integral part of modern and industrialised societies [1], often considered a cultural advantage. The term equity can be better understood in comparison with the term equality, for which different declinations exist depending on the application context. Here, we focus on the educational field. As argued by de Los Santos et al. [2], equality entails ensuring that every individual's needs are met by providing them with equal access to resources and opportunities. On the other hand, equity concerns a system that redistributes common resources to establish institutions and educational facilities with a higher probability of achieving greater equality. Social inclusion is also among the desired results of equity [3], [4], [5], [6]. Inclusion aims to create a united and cohesive environment, where individuals are highly integrated [7]. The European Union has studied the effect of exclusion/inclusion of individuals in society for several years and promotes actions to improve social inclusion [8], [9], [10], [11], [12], [13], [14].

The social inclusion problem is widely considered in education [15], [16]. Lack of inclusion in primary schools can have long-lasting effects on children's physical, social, and psychological development. When children do not feel included in their school environment, they can develop problems such as low self-esteem, anxiety, and depression [17]. Additionally, exclusion from social activities and peer groups can exacerbate the issue, resulting in poor academic performance and limited opportunities for success later in life [7], [18]. Thus, it is crucial to address this problem early on by providing an inclusive and supportive environment where all children feel welcome and valued [19].

Social inclusion is typically assessed through surveys, direct observation, or third-party recounts of social interactions. However, these traditional approaches have substantial limitations, including subjectivity, cost inefficiency, and privacy concerns, which complicate their use, especially in dynamic and sensitive environments like schools [20]. Recent advancements in Internet-of-Things (IoT) technologies and network analysis provide promising alternatives by offering objective, real-time, and scalable methods to capture social interactions [21]. IoT devices, particularly those based on Bluetooth Low Energy (BLE), have been increasingly used in studies of social interactions due to their ability to provide accurate proximity detection and interaction tracking. Network analysis further complements IoT data collection by quantifying social structures and identifying isolated individuals through well-established metrics and measures. Leveraging this synergy, our work proposes an innovative approach integrating IoT-based proximity detection with social network analysis techniques to quantitatively identify and address social isolation among individuals in educational settings.

To the best of our knowledge, this is the first work that proposes the usage of IoT devices for the measurement of social inclusion of individuals, with the suggestion of interventions to improve the inclusion of the less integrated group members.

A Framework for Social Inclusion

The above premises motivate our work, which is a framework for measuring social relationships and increasing social inclusion. In a nutshell, we aim to introduce an approach based on the usage of IoT devices for the detection of the relationship network among students and on the application of social network analysis to identify students at risk of isolation and support educational initiatives aimed at increasing their inclusion. In the following, we describe the way we intend to measure the level of isolation, which essentially depends on the frequency and duration of the contacts a student has. To make our approach feasible, we design and use affordable over-the-counter IoT devices to detect social contacts. We are aware that reducing isolation is not a sufficient condition to ensure social inclusion, but it is one of the necessary factors for this purpose, and therefore, it is worthwhile to consider actions aimed at its reduction. In the remainder, we structure the presentation of the framework between its theoretical and practical components.

In Section 2, we first present the assumptions from social theory we start from, followed by the definition of the mathematical model we use to represent the network under study and the measures used to quantify the degree of integration of the network's members. This part delves into social theories and human behaviour, providing a basis for developing measuring methods for social relations using network analysis. The basic idea is to characterise the position of students in the community they are part of in their school settings. These measures allow us to identify the less integrated students and suggest policies to increase their integration. The policies take into account the position of the other students in the network as well as the natural tendency of the students in the network to establish relationships with one another. In particular, the policies proposed by our framework balance the increase in inclusion with said tendency. The rationale is to propose policies that minimise the inherent opposition the network would pose – due to the endogenous forces that characterise the observed emerging behaviours – against the exogenous enforcement of the suggested inclusion intervention. We conclude Section 2 with the definition of an algorithm that proposes policies for increasing the inclusion of the less integrated individuals. These policies indicate the establishment of new connections (and their expected strength in terms of frequency and duration) between previously unconnected nodes in the network. Then, domain experts (e.g., sociologists) can concretise the proposed intervention by determining their nature (e.g., doing some specific activities), frequency, and duration.

The practical part, found in Section 3, defines technologies, algorithms, and a pipeline leading from the tracing of interactions to the reconstruction of the network of relationships among the studied individuals. In particular, we base our framework on the definition and construction of prototype affordable over-the-counter IoT devices that the studied subjects wear and that detect their social interaction in a fully distributed and range-free manner – i.e., without requiring any central device or network component to keep the recording units connected. Using these low-cost devices not only lowers the barriers to entry for researchers, but they make the collection of data feasible when other methods are not affordable/available. Alternatives include the conduction of surveys to query the participants about their interactions, and third-party recounts of observed interactions, e.g., done by the researchers live or via recordings. While it might be difficult to gather valid data from surveys when studying children, as can be the case in the context of education, the observation of the participants is a viable option. However, this option is costly, privacy-threatening, and laborious if done with video-recording equipment or time-consuming and more prone to errors if done by live observers. The usage of IoT devices for the detection of social interactions is not new in the field [22], [23], [24]. Compared to the alternatives, deploying our experiments just entails the deployment of the devices themselves, while the alternatives usually require the installation of additional equipment, like base stations and routers, which can additionally impose restrictions on the behaviour of participants, e.g., because they restrict the area where the participants can interact. In Section 3, we define the methodology and data pipeline and describe the salient details of the construction, programming, testing, and initial validation of the IoT devices. In particular, we specify a multi-stage algorithm for the detection of social interactions that starts from the collection of beacon-strength signals among the devices (through empirical measurements) and produces the aggregated measure of connection strength (in terms of overall frequency and duration) among the devices/subjects.

After having defined the theory and practice of our framework, in Section 4, we start validating it via a real-world case study of an elementary school class in Bologna county, Emilia-Romagna region, Italy. At the meta-level, the case study confirms the suitability of our approach.

Overall, the structure of the article includes Section 2, where we introduce the theoretical foundations and models for measuring social inclusion, Section 3, where we detail our IoT-based methodology and data processing approach, and Section 4, in which we validate the framework through a primary school case study. Sections 5 and 6 close the article by respectively positioning our contribution and discussing final remarks and future directions.

2 Sociological Concepts and Network Analysis

We now introduce the theoretical foundations behind our work. There are two main aspects to consider: the sociological theory underlying our interpersonal relationship model and the scientific aspects of network theory and analysis used for modelling such theory. Specifically, we rely on sociology to formally describe what it means to detect the presence of face-to-face interactions and to infer relationships among people, while we use network theory to explain how we can measure said relationships between the entities in the network.

2.1 Contacts, Interactions, and Relationships

In social science, the description of relationships within a social group has been done using the tools of network analysis, such as the sociogram developed by Moreno [25], which is a graphical representation of the connections between people built using questionnaires. By social group, we mean a set of individuals who interact according to certain patterns, who have a sense of belonging to the group, and who other members consider as part of the group [26]. In our case, the social group under consideration is the classroom and the relationships are those between the students. Defining social relations is a crucial step in our study and we build on the social relation structure described in Sztompka's hierarchy of social behaviours, also called "sociological hierarchy" [27]. Following Sztompka's framework, we rely on three hierarchical elements: social contact, social interaction, and social relation. Social contact refers to the instances when individuals find themselves in a situation where they meet. These kinds of situations can include any type of encounter or connection between people, whether intentional or incidental. Social interaction occurs when two or more individuals are involved in face-to-face interaction and in a context where they act with reference to each another [28, p.304]. This process may involve verbal exchanges, gestures, facial expressions or any other form of verbal or non-verbal communication. A necessary condition for social interaction detection involves prolonged social contact over time. A social relation is a more enduring connection between individuals that develops over time. These kinds of connections can include family bonds, friendships, professional relations, or any other type of connection that persists over time. In Sztompka's model, social relations are the sum of the social interactions between individuals throughout time [29, ch.22].

Since we aim to provide a measure of social inclusion, we need to record individuals' social interactions, which, according to our framework, are the building blocks of social relationships. Hence, we need a criterion to define relationships from interactions and contacts. From Sztompka's model, social interaction is the sum of social contact lasting over time. In this work, we also draw inspiration from Behaviour Analysis [30] and define the relationship between two individuals given the *frequency* and *duration* of their social interactions, and we identify a social interaction when a social contact lasts more than a given time threshold.

However, there are no reference values for such a *contact threshold* in the literature – i.e., the minimum duration of social contacts to become interactions – because it strongly depends on the context in which these actions take place. Indeed, the duration of the interactions can change depending on the context of the experiment, e.g., if the agents are students, and we run the experiment in different dynamic contexts (e.g., playground vs classroom activities). Since the

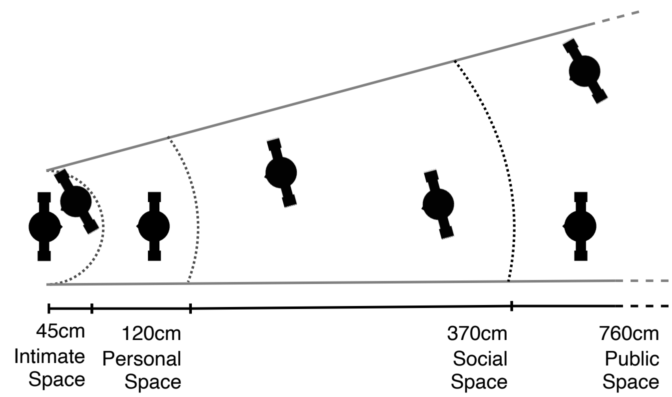


Fig. 1: Schematisation of proxemic distances.

definition of the contact threshold is context-dependent, we do not fix a general value but rather adapt it to both the conditions of the experiment and the empirical data obtained from the devices – which might require us to adjust the value depending on the delay they impose on the reception of signals (cf. Section 3).

Besides timings, we also need to establish a measure of distance for the detection of contacts. Specifically, we want to identify face-to-face social interactions, which compels the definition of social interactions, face-to-face, and social distances:

- *Social interaction* is the process of reciprocal influence exercised by individuals over one another during social encounters. It involves verbal and non-verbal communication and can occur in various forms, like vis-à-vis conversations and group discussions [31].
- *Face-to-face interaction* is a social interaction carried out without any mediating technology [32], and it corresponds to the mutual influence of individuals' direct physical presence with their body and verbal language [33, p.15], [34, p.50]. Face-to-face interactions allow people to communicate directly. Having social interactions and being part of a social group can improve mental health, for example by reducing depression and anxiety [35, p.15].
- *Social interaction distances* were first studied and defined by anthropologist Edward T. Hall in the 1960s. Hall's coined the term "proxemics", which theorises how people use space during social interactions in different cultures [36]. His research identified four types of social interaction distances: intimate, personal, social, and public distance, which are still widely used today. The four areas based on social interaction distances are represented in Fig. 1.

Given the definitions above, we consider face-to-face interactions valid if they occur within the 120 cm personal distance between the two subjects' fields of view. Therefore, we are considering the intimate and the personal spaces. This distance is determined according to the human field of view, which considers the direction of the upper body as the view's central point using the binocular range of 120 degrees, granting a comfortable interaction. We deem this assumption reasonable in the context of a group, where the primary interactions are in-person, and we aim to value the ones that occur intentionally, involving both verbal and non-verbal levels, including body language and visual contact.

2.1.1 Network Analysis to Measure Social Inclusion

Now that we defined the relationships we want to study from the sociological standpoint, we move to introduce the structure that we use to represent these relations and the set of measures useful for gaining insight and direct interventions on the network they constitute. We develop this step following a Network Analysis approach.

Network's structure

We model the network as a graph $G = (V, E)$. V is the set of n nodes $v_1, \cdots, v_n$ in the network (e.g., the students in a given class). Since the theory we consider defines relationships based on the measured social interactions, with their duration and frequency, we represent relationships between nodes with two-dimensional irreflexive undirected weighted edges, i.e., each edge $e \in E$ is a tuple $e = (v, v', t, f)$ that associates the pair of nodes it links (v and v') with two weights. The first, t, is the *duration*, and it corresponds to the sum of the periods of each interaction between v and v' within the data-collection phase (i.e., the cumulative time v and v' interacted); the second, f, is the *frequency*, and it corresponds to the number of interactions v and v' had during the data-collection phase. We detail the process of interaction detection in our framework via IoT devices in Section 3.

2.2 Network Measures for Estimating Social Isolation

We introduce a set of measures and a method to aggregate them to quantitatively define and identify isolated agents in the studied network.

2.2.1 Measuring Social Inclusion

We start by introducing two complementary network measures for estimating the degree of social inclusion perceived by the agents.

PageRank Centrality

The first measure for estimating how integrated a node is in a network regards looking at the number of direct connections it has with the other nodes, given the latter level of integration. The idea of basing the measure of the centrality of a node on the centrality of its neighbours originally comes from Eigenvector centrality [37, ch.7.1.2].

Eigenvector centrality measures the importance of a node in a network by assigning higher scores to those connected to other high-scoring nodes. While it seems reasonable to consider a node integrated if it has connections with other integrated nodes, we think that Eigenvector centrality is not apt for measuring social inclusion. Indeed, the measure also assigns high scores to nodes with only a few high-scoring connections, e.g., we could have high-scoring nodes with only one high-scoring connection, which hardly seems an acceptable definition of integration.

PageRank [38] is a variant of Eigenvector centrality whereby the centrality bestowed by a node to its neighbours is evenly divided among them – more concretely, PageRank refines Eigenvector centrality by discounting the centrality passed by a neighbour to a node by that neighbour's degree. While this mechanism was originally introduced to estimate the importance of web pages (so that a central page, e.g., a directory, that leads to many pages bestows little centrality to the latter) we deem it useful to measure an appropriate definition of integration – where (well-integrated) nodes with connections to many neighbours pass only a small amount of centrality to each of the latter. Essentially, using PageRank means that, although a node v has an interaction with a well-integrated node v', if v' has many interactions with also other nodes, v will receive a fraction of the "integration" that v' enjoys – seen at a person's level, it means that v' has to divide their attention among their many neighbours of which v is part of. We believe that PageRank represents an appropriate interpretation of integration also thanks to studies, such as by Dunbar and others [39], that demonstrate the limits on the number of meaningful relationships that an individual can simultaneously sustain, beyond which their relations become shallower and fragmented – i.e., eroding the feeling of integration they bestow to their connections.

Since connections among agents have different durations and frequencies, we give a pondered interpretation of PageRank where each neighbour contributes to a node's score considering a function of its edge's (e_i) duration (t_i) and frequency (f_i). The motivation for abstracting the weighing of the edge is that, depending on different contexts, researches can consider the usage of different formulas that compound frequency and duration – possibly also discarding one dimension to solely focus on the other.

Summarising, by using PageRank centrality we introduce a direct-connection measure that both considers the topology of the network that surrounds a node and the number of connections each node has, pondered by the duration and frequency of their interactions. Let $k_{v'}$ be the degree of node v', i.e., the number of edges connected to it and $\omega(t, f)$ an edge-weighing function parametrised over duration (t) and frequency (f). Formally, we define PageRank centrality pr as in Equation (1).

$$pr(v) = \sum_{(v,v',t,f) \,\in\, E} \frac{pr(v')\ \omega(t,f)}{k_{v'}} \tag{1}$$

In Equation (1), the score for node v considers the scores $pr(v')$ of all nodes v' connected to node v, discounting each contribution by its degree $k_{v'}$, while also incorporating the weight of the edge estimated through ω, considering the edge's duration (t) and the frequency of interactions (f).

We can resolve Equation (1) by reformulating it in matrix form to leverage eigenvalue decomposition. In that form, the PageRank vector corresponds to the leading eigenvector of the system of linear equations induced by each node's centrality formula. Convergence of the computation is ensured by the fact that the centrality scores are non-negative so that we can apply the Perron-Frobenius theorem, whereby a matrix that has only non-negative elements has only one eigenvector with only non-negative elements, which corresponds to the leading eigenvector [40], [41].

In this study, we choose to instantiate the edge-weighing function as $\omega(t, f) = t/f$, i.e., where duration contributes positively to the score a neighbour supplies to a node while frequency detracts from that score by dividing the duration value. This interpretation stands on the principle of favouring longer interaction times, which more likely allow individuals to bond and interact in meaningful ways, rather than having more rapid interactions that allow only superficial relationships. The proposal for this simple yet relevant instantiation of function ω comes from previous work that evidence the link between the duration and frequency of the social interaction in the perceived degree of integration/isolation of subjects [42], [43]. The choice of using the t/f ratio over weighing edges, e.g., based solely on duration or frequency, stems from such empirical evidence. Indeed, duration alone fails to capture the consistency of social engagement because it corresponds to the sum of all recorded interactions between the agents, which lacks the dimensionality of frequency (i.e., how many interactions make up the total duration). Conversely, frequency alone may overemphasise numerous brief encounters that lack the depth required for meaningful social bonding. Nodes exhibiting numerous brief interactions may appear well-connected under frequency-based metrics, yet they can potentially feel social isolation when these interactions lack meaningful duration. Thus, the ratio t/f provides a composite measure that rewards sustained, meaningful interactions while penalising fragmented, superficial contact patterns. We deem this distinction (quantified by the t/f ratio) crucial for applications in social network analysis where the focus of the study is understanding social integration, rather than mere contact patterns. This approach aligns with sociological theories suggesting that social integration depends not merely on contact frequency, but on the quality and depth of interactions, which correlate positively with the duration of interactions [44], [45].

Summarising, the $\omega(t, f)$ instantiation we propose influences the resulting PageRank scores by redistributing importance towards nodes engaged in "substantial" interactions. Therefore, nodes participating in extended, meaningful exchanges receive high centrality scores, even if their overall connection count remains modest. In particular, this instantiation identifies individuals who may appear peripherally connected in terms of raw contact but maintain deep, influential relationships within the network. We emphasise that, while in this study we establish a linear inverse correspondence between duration and frequency, the proposed PageRank measure is parametric to one such instantiation, allowing it to capture other kinds of phenomena and adapt to different contexts and domains that require alternative interpretations of the duration and frequency parameters to provide a measurement of the recorded interactions that align with specific contexts – e.g., in situations where duration and frequency lie within constrained ranges and require normalisation. Other examples of instantiations of ω could use the ratio of the logarithms of duration and frequency to highlight differences in magnitude or introduce a weighted combination such as $\omega(t, f) = \alpha\, t/f + (1-\alpha)\, \log t/\log f$ to balance absolute and relative effects (adjusted through variable α), or employ a power-law transformation, like $\omega(t, f) = t^{\beta}/f^{\gamma}$ to accentuate either sustained interactions (greater β values) or repeated encounters (greater γ values) depending on the chosen exponents. In this instance, we prefer a simple definition of the ratio and leave the study of contextualising alternatives to future work.

Since we want to compare the strength of a given integration policy across different measures (as discussed in Section 2.2.2), we use a normalised form of pr, called npr, defined in Equation (2) and based on the marginals of the overall sum of the PageRank centralities.

$$npr(v) = \frac{pr(v)}{\sum\limits_{v' \in V} pr(v')} \tag{2}$$

Betweenness Centrality

The second measure we use to detect isolated individuals is Betweenness centrality, which identifies isolated nodes as those that rarely lie on the shortest paths that connect other nodes in the network [37, ch.7.1.7]. The interpretation we give to Betweenness is that nodes that work as mediators for other nodes are more likely to take part in group social interactions and, thus, be more integrated. Formally, since multiple shortest paths can exist between any given pair of nodes in a network, we calculate the Betweenness of a node by summing the number of the shortest paths between all pairs of nodes in the network that pass through that node, divided by the total number of the shortest paths between the pairs of nodes in the network. Let $sp(v_1, v_2)$ return the number of the shortest paths from v_1 to v_2 in a given network and $sp(v_1, v_2, v_3)$ return the number of the shortest paths from v_1 to v_2 that pass by v_3 in the network. Formally, we define Betweenness centrality bw in Equation (3) and its marginals-based normalised form, called nbw, in Equation (4).

$$bw(v) = \sum_{(v_i, v_j, t, f) \in E} \frac{sp(v_i, v_j, v)}{sp(v_i, v_j)} \tag{3}$$

$$nbw(v) = \frac{bw(v)}{\sum\limits_{v' \in V} bw(v')} \tag{4}$$

Betweenness makes a different aspect of integration/isolation emerge than the one we focussed on using PageRank. Indeed, notice that the measure that we define is oblivious to the weights of the edges connecting the nodes. If we weighted the measure wrt the frequency f, we would rank nodes with more frequent interaction higher (irrespective of the time spent on each one), i.e., we would see as better integrated those nodes that have chances to see its neighbours multiple times. On the contrary, if we weighted the measure wrt the duration t, we would rank higher nodes that "spend" (irrespective of the frequency) more time interacting with their neighbours. Of course, one could alternatively want to give a higher value to "meaningful" relationships, marked by higher ratios between duration and frequency. Here, we purposefully choose to make the measure weight-oblivious to mark that nodes with high betweenness tend to work as connectors between relatively separated parts of the network. Thus, we view them as integrated into the network since information (about activities, gossip, etc.) likely passes by them to reach different network clusters – following evidence from studies that link gossip networks to social integration [46], [47].

On the Complementarity of PageRank and Betweenness

Since we propose PageRank and Betweenness as complementary measures for social inclusion, we conduct a small simulation-based study on small-world networks [48] to investigate the extent to which Betweenness and PageRank centralities provide complementary perspectives on node scores. We select small-world networks because they capture key structural features commonly observed in real-world social networks like classrooms, workplaces, and other social environments where individuals are often linked through denser clusters while still maintaining occasional intra-cluster connections.

Note that this study only gives an indication of the correlation between these two measures in the general case (i.e., as defined in the literature) and not the specific ones we propose above (in particular, the PageRank variant) since we do not have statistically-relevant instances of edge weights (duration and frequency) to support realistic simulations.

We ran 100 independent small-world network realisations considering networks of ca. 50 nodes (st. dev. 2.8) and an average degree of 6.40 (st. dev. 1.65). We calculate both the Pearson and Spearman correlation coefficients between the two measures, obtaining respectively an average of 0.588 (0.093 st. dev.), and 0.540 (0.105 st. dev.). Although the correlation is positive, it is weak, indicating how PageRank and Betweenness can cover different interpretations of integration in the kind of networks under study.

Our results are aligned with those reported by Yin, et al. [49], who study (among others) the correlation between the two measures under different kinds of networks.

Estimating Resistance to Integration Policies

The intervention to integrate more isolated nodes in the network is an exogenous force that might be counteracted by the endogenous ones at play within the network, i.e., the factors that influence the tendency of the nodes to establish relations. Hence, we try to increase the chances that the network would sustainably incorporate the exogenous actions for integration by prioritising those interventions that minimise the deviation from the measured tendency of the nodes to establish connections. We favour this policy to avoid imposing stronger alterations that would both require more energy (e.g., impositions and reiterated interference) and could be more prone to failure (e.g., due to the aversion of the involved subjects, who might perceive the intervention as unreasonably coercive).

Mathematically, we interpret this trait as the tendency of nodes to connect with other nodes, given their respective degrees. Of course, other attributes can provide alternative views on the tendencies of the network, e.g., one could study the tendency of nodes to connect to other nodes seen as integrated (given some definition of integration, e.g., using npr and nbw defined above). Since we want this measure to be as orthogonal as possible wrt the concepts of integration defined above, we choose the generic measure of assortativity by degree, which simply tracks the tendency of nodes to establish connections with other nodes that either have an equal, lower or higher amount of connected nodes.

Assortativity Coefficient by Degree

Formally, the assortativity coefficient by degree configures as a Pearson correlation coefficient between the degrees of connected nodes in the network [50]. Formally, we define the auxiliary functions ξ_{ij} and δ_{ij} as in resp. Equation (5) and Equation (6).

$$\xi_{ij} = \begin{cases} 1 & \text{if } (v_i, v_j, t, c) \in E \\ 0 & \text{otherwise} \end{cases} \tag{5}$$

$$\delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases} \tag{6}$$

In essence, ξ_{ij} is as an indicator function for edge existence, while δ_{ij} represents the Kronecker delta for node identity.

Given the definitions in Equations (5) and (6), and let k_i be the degree of node v_i, we can define the assortativity coefficient by degree r as reported in Equation (7).

$$r = \frac{\sum\limits_{v_i, v_j \in V} \left(\xi_{ij} - \frac{k_i\ k_j}{2|E|} \right) k_i\ k_j}{\sum\limits_{v_i, v_j \in V} (k_i\ \delta_{ij} - \frac{k_i\ k_j}{2|E|})\ k_i\ k_j} \tag{7}$$

The key insight underlying Equation (7) is that the assortativity coefficient measures the correlation between node degrees across edges, normalised by the expected correlation under a random model. The numerator captures the covariance between degrees of connected nodes, while the denominator represents the variance in the degree distribution. The term $(k_i\ k_j)/(2|E|)$ in both the numerator and denominator corresponds to the expected correlation under the configuration model, where the edges are randomly rewired while preserving the degree distribution.

The measure ranges from -1 to 1. Positive values indicate assortative mixing (where high-degree nodes preferentially connect to other high-degree nodes, forming hierarchical, multi-layered structures), negative values suggest disassortative mixing (high-degree nodes tend to link to low-degree nodes, forming sparse star-like formations), while values close to zero imply random mixing.

2.2.2 A Procedure to Improve Social Inclusion

We can now define a procedure to improve the social inclusion of more isolated individuals. Essentially, the procedure identifies the least integrated node v in the network and proposes the establishment of a new connection with a node v' which v does not have a connection to such that a) v and v' have a common neighbour that can work as "bridge" and b) the establishment of the connection minimises the deviation from the natural tendency of nodes in the network to connect. The procedure also provides an expected weight (in terms of frequency and duration) for the newly established relation. From the sociological perspective, we see this intervention suggestion at the base of an integration policy for a domain expert to act upon. In particular, the suggested intervention builds a new "desired" configuration of the network after the application of the policy.

Since this is an important point of the framework, we further clarify it with an example. Imagine, as we do in Section 4, that we have a network of relationships within a class of students. We find v to be the least integrated one and, running our procedure, we have the suggestion to connect v to v', mediated by v'', with a certain frequency and duration. Given this intervention proposal, a psychologist defines a policy that, within a set period, sees v, v', and v'' spend time together, e.g., playing and doing homework. After the application of the policy, we can repeat the experiment and test whether *i*) the connection between v and v' exists and *ii*) how close its weight approximates the expected network configuration, giving a way for the psychologist to track and adjust the policy wrt the effects it had on the network. In general, we see the definition of the policies that implement the suggestion from the framework orthogonal to the latter. We discuss the matter in a dedicated part of Section 6 on policy suggestions.

Social Inclusion: Measurement and Improvement

To concretise the ideas above, we start by introducing the measure $\iota(v)$, which is the inclusion index used to provide an overall indication of the inclusion of nodes. Mathematically, as reported in Equation (8), the index corresponds to the maximal value between the normalised PageRank (Equation (2)) and Betweenness (Equation (4)) centralities.

$$\iota(v) = \max\left(npr(v), nbw(v)\right) \tag{8}$$

Since we compute and evaluate the integration policies over possible what-if scenarios where we connect isolated nodes with others they do not have connection to, we introduce a version of the assortativity coefficient by degree r that is parametric to the network $G' = (V', E')$ it is computed on, i.e., the E and V in the definition of r for $r(G')$ are respectively V' and E' – and, thus, the fixed definition of r given earlier corresponds to $r(G)$. We do the same for ι, where $\iota(G, v) = \iota(v)$ and $\iota(G', v)$ computes the values of npr and nbw on the edges and nodes found in G'.

The steps we propose for defining an integration policy for the most isolated individual of G are:

1) Identify the most isolated node, which is $v \in V$ s.t. $\iota(v)$ is minimal (picking one at random if more than one qualify).
2) Compute the possible what-if scenarios where we connect v to a node $v' \in V$ it has no connection to but there is a 2-step path between v and v' mediated by a common neighbour v''. Formally, we find the set of candidates W as defined in Equation (9) (the formula omits durations and frequencies from the edges with $\cdot$ since they are immaterial for the selection).

$$W = V \setminus \left(\begin{array}{c} \{v' \mid (v, v', \cdot, \cdot) \in E\} \,\cup \{v' \mid \nexists v'' \mid \\ \{(v, v'', \cdot, \cdot), (v', v'', \cdot, \cdot)\} \subseteq E\} \end{array} \right) \tag{9}$$

W contains the nodes v' not connected to v that have a common neighbour v''. Note that, to find possible new connections for the most isolated node v, we assume that v is connected to at least one neighbour[1]. The last ingredient to build our what-if scenarios is the specification of the duration t and frequency f of the connection between v and the nodes in W. To avoid imposing on v interaction regimes outside their empirically-determined predisposition, we first select which neighbour shall mediate the interaction between v and v' and then use the minima of the duration and frequency of the respective v-v'' and v'-v'' edges as resp. the duration and frequency of the new v-v' arc – i.e., we set the duration and frequency of the new arc to the maximum "allowed" by the resp. relationships of v and v' with v'', which is determined by the minimal between each pair of values. Formally, we identify the set of candidate mediators or "bridges" between v and v' using function β, defined in Equation (10).

$$\beta(v, v', E) = \\ \left\{ (v'', t'', f'') \;\middle|\; \begin{array}{l} \{(v, v'', t, f), (v', v'', t', f')\} \subseteq E \\ \wedge\, t'' = min(t, t') \wedge f'' = min(f, f') \end{array} \right\} \tag{10}$$

If more than one neighbour qualifies according to this measure, we select one of them randomly. The measure selects the neighbour that maximises the time v and v' can spend together while minimising the frequency of their interactions (we recall that, in this work, we prefer longer interaction periods among the nodes, on the ground that longer periods more likely support the establishment of significant relationships). Hence, we use t'' and f'' as the resp. time and frequency values of the new connection between v and v'. For each $v' \in W$, we build a new network $G' = (V, E \cup (v, v', f'', t''))$ which is a variant of the original network G where v is connected to v' ("through" a common neighbour v'') with duration t'' and frequency f''.
3) Now that we have the what-if scenarios, we define the selection logic among them so that we find the one that strikes a balance between increasing the integration index produced by the inclusion of the new edge and keeping small the deviation from the measured tendency of the nodes to establish relationships (measured by the assortativity coefficient by degree). Mathematically, we select G' such that it maximises the measure, as defined in Equation (11).

$$\left(\iota(G', v) - \iota(G, v)\right)\left(1 - ||r(G')| - |r(G)||\right) \tag{11}$$

In Equation (11), we use ι to quantify the integration gain of v from G to G'. Since this change may or may not be in line with the natural tendency of the network to form links between nodes, we discount the gain given by G' wrt the 1-complement of the deviation of r computed on G', i.e., the more the deviation approximates zero, the more the gain given by G' increases. The absolute value of the difference of the absolute values of the measure on the two networks allows us to include both negative $[-1, 0)$ and positive correlations $(0, 1]$.

On Choosing Two-Step Inclusion Interventions

In the definition of our algorithm for intervention suggestion, we restrict intervention candidates to two-hop neighbours. The reason for such a restriction is related to policy implementation feasibility.

On the one hand, one could completely dismiss the need for multi-hop policies and directly connect an isolated node to any non-neighbour node that satisfies the constraints

1. Besides the formal justification behind this assumption, we morally rationalise the assumption by considering the case of totally isolated individuals, who require policies outside the scope of our framework.

of our formulation (of course, minus the two-hop distance requirement and related edge weight estimation). However, one such intervention can be impractical to implement since the two agents do not normally interact with each other, i.e., the agents could develop high opposition to an extraneous intervention that they perceive as evidently forced. For this reason, we refine our suggestion algorithm to consider multi-hop suggestions where in-between actors mediate the constitution of a new connection between the isolated agent and the one that increases the former's integration.

Considering multi-hop strategies, two-hop interventions represent the minimal energy configuration that can bridge isolated network components while requiring coordination among the fewest number of individuals. Indeed, each additional hop in the intervention path introduces a growing set of constraints that the social operator must consider when crafting implementable policies, as every intermediate node represents a stakeholder whose agreement and compliance become necessary for successful intervention deployment. Extending beyond two hops would theoretically enable the discovery of intervention strategies with potentially superior network integration properties. However, such extended paths suffer from severe scalability limitations in practical implementation scenarios – requiring the social operator to negotiate and maintain compliance across longer chains of intermediary actors, each with their own incentives, constraints, and potential points of failure.

For these reasons, we focus on minimal two-hop policies to balance intervention effectiveness and practical implementability. While we fix the two-hop policy in this study, we deem refinements promising directions for future research. Indeed, more sophisticated incarnations of the suggestion algorithm could explore redundant two-hop combinations, where multiple parallel intervention paths provide robustness against individual agent non-compliance. Additionally, a parametric version of our approach could dynamically adjust the hop distance based on network characteristics, intervention success rates, or available coordination resources, enabling adaptive intervention strategies that optimise for specific deployment contexts while catering to deployment feasibility.

Edge Cases and Multi-Step Intervention Scenarios

Notice that, following our formulation of the suggestion algorithm, it is not guaranteed that we can find a G' such that ι is positive for the given v. As an example, consider the small network shown on the left-most panel of Fig. 2 (we use a small-scale example to provide a succinct description, but the phenomenon is scale-independent). In all panels of Fig. 2, we parametrise the size of the nodes wrt their ι and the size of the edges wrt the ratio duration over frequency. We report the weights directly on the edges in the figure – in the format (t, f) – while the diameter of a node corresponds to its ι. In the left-most panel, the node labelled 2 has the lowest ι – as visible from its size. We compute the possible alternative interventions and visualise them in the remaining panels of Fig. 2. The alternatives, in left-to-right order, add an edge between 2 and resp. 0 (through 3) and 1 (through 4). In all cases, the score for the alternative network is negative – indeed, 2's ι (and its size in the panels) decreases by a small amount in the alternatives.

While these edge cases are a consequence of the mathematical definition of ι and the scoring measure of the interventions, we conjecture that the measures (given the principles we followed to define them) reflect edge-case social conditions that require more complex actions than the relatively straightforward ones suggested by our proposal. For instance, one solution to improve the inclusion of 2 in the example is to first connect 1 and 3, decreasing the strong coupling between 0 and 1, and then introduce the connection for 2. The exploration of this multidimensional space of solutions is vast. Indeed, while one could argue for proposing practical solution of 2 or 3 steps (e.g., taking the second most isolated node, etc.), we only deem it appropriate to explore the problem from a parametric viewpoint (i.e., considering the n-dimensional solution that strikes the balance between inclusion and network change). However, we see one such study outside the scope of this work – and, thus, for the definition of our procedure, we consider the edge cases of non-positive scoring outside the domain of solutions of the framework.

In general, besides following a multi-step approach based on adding connections, more complex intervention strategies can also suggest steps like establishing other connections or decreasing the strength of the existing ones before proceeding with the introduction of the new relation for the isolated node. A concrete example of the need for these advanced strategies is when the strength of the connections among the members of a tightly connected cluster prevents "outsiders" from establishing meaningful connections with them [51], [52]. As stated, since the aim of this article is to provide an overall presentation of our framework for studying social inclusion, we leave the exploration of more advanced and complex strategies for future work.[2]

Computational Complexity and Convergence of the Intervention Algorithm

We close this section by briefly analysing the computational complexity and convergence of the proposed algorithm.

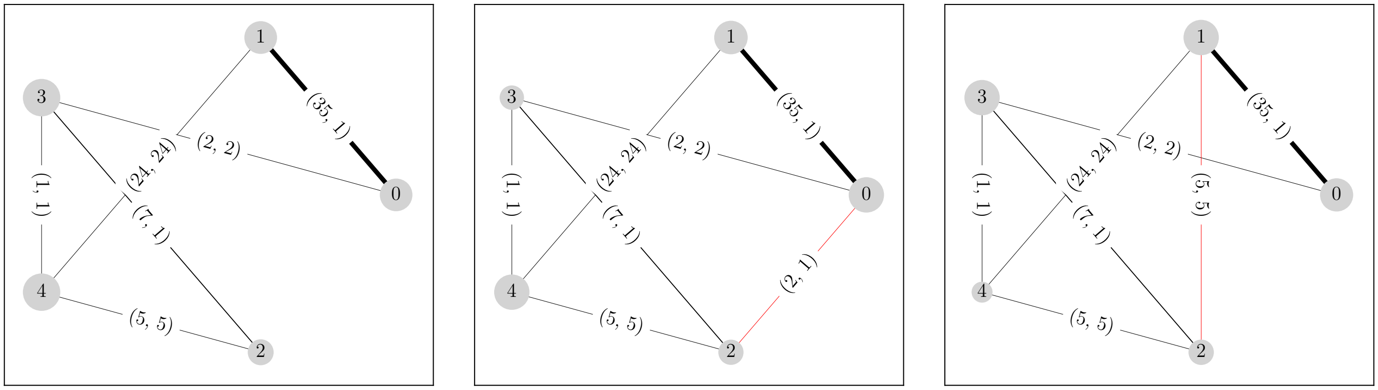
We begin with the computational complexity analysis, starting from the separate measures used by the algorithm.

The PageRank calculation constitutes one of the most computationally intensive components. Indeed, to measure PageRank we have to solve the eigenvector equation. Using matrix multiplication-based eigenvector techniques, the complexity reaches $O(|V|^3)$, however there exist approximate, iterative methods typically requiring $O(m \cdot |E|)$ operations, where m represents the number of iterations needed for convergence. In practice, $m \in [10, 100]$ iterations are typically sufficient for acceptable approximations.

Betweenness centrality presents a complexity profile of $O(|V| \cdot |E|)$ using Brandes' state-of-the-art algorithm [54], while the complexity of the assortativity coefficient by degree is $O(|V|^2)$.

Calculating ι implies computing both PageRank and Betweenness, resulting in $O(m \cdot |E| + |V| \cdot |E|)$, which we calculate for all nodes, amounting to $O(|V| \cdot (m \cdot |E| + |V| \cdot |E|))$

2. Multi-step procedures that involve solution-space exploration have a high computational complexity. State-of-the-art solutions for these problems employ advanced approaches from operations research and constraint satisfaction techniques [53], whose adaptation to our case deserves dedicated development.



Starting network | Alt. 1: 2 connected to 0 (br. 3) | Alt. 2: 2 connected to 1 (br. 4)

Fig. 2: Example of edge case of only negative score solutions.

Moving on, following the algorithm for suggestion, the connection selection phase exhibits $O(|V|^2)$ complexity in the worst case. Indeed, for each potential connection candidate v', the algorithm must identify common neighbours through set intersection operations, which requires examining all neighbours of the node under scrutiny v (with degree k_v) against all neighbours of each candidate v' (with degree $k_{v'}$), resulting in $O(k_v \cdot k_{v'})$ operations per candidate pair.

The subsequent evaluation requires computing ι and assortativity for each candidate connection.

Combining all phases, the algorithm exhibits complexity $O((V+|V|^2)\cdot(m\cdot|E|+|V|\cdot|E|))$, or $O(|V|^2\cdot(m\cdot|E|+|V|\cdot|E|))$ for simplicity. We note that this upper-bound complexity regards dense graph structures where most nodes have potential two-hop connections to the isolated node. For sparse graphs, which commonly occur in social networks, the actual complexity reduces significantly – social networks typically exhibit small average degrees (often $\propto \log(|V|)$), reducing complexity in practical instances.

Moving to convergence, PageRank benefits from strong theoretical convergence guarantees rooted in the Perron-Frobenius theorem. Similarity, Brandes' algorithm [54] for Betweenness and the definition of assortativity guarantee the convergence of both calculations.

The intervention selection algorithm has an objective function (Equation (11)) that combines ι with r (assortativity), which seems to entail a multi-objective optimisation problem. However, we define a greedy approach for selecting a single-step "best" intervention, without considering multiple intervention rounds[3]. This greedy approach guarantees convergence, although, as argued when discussing edge cases, it can amount to no additional connections able improve the objective function.

3 Framework Definition and Device Prototyping

As described in the previous sections, our framework is grounded in the theory that social inclusion can be derived through social interactions, with each interaction being a composite of face-to-face contacts. This conceptualisation forms the basis of our practical approach, enabling us to measure social relations in a structured and quantifiable manner. In this section, we detail the framework methodology behind our proposal and the prototype developed in this work. In particular, we illustrate how our innovative approach not only operationalises the theory of social inclusion through interactions but also addresses existing challenges in measuring and enhancing social relations within educational settings.

3.1 Methodology

The architecture of our framework, as illustrated in Fig. 3, spans from the initial face-to-face data gathering to the comprehensive network analysis of social interactions. The proposed framework is designed as a sequential data chain, encompassing the following stages:

- *Data Collection*: Gather raw data on social interactions, primarily through a distributed network of IoT devices that track face-to-face contacts.
- *Data Processing*: Convert the raw data into a structured format suitable for analysis. This stage involves steps such as data cleaning, integration, and transformation.
- *Data Analysis and Visualisation*: Analyse the processed data to extract meaningful insights about social interactions and relationships. This stage also involves visualising the data to make the findings more accessible and understandable.
- *Evaluation and Reporting*: Assess the quality and effectiveness of social relations and the nature and impact of any inclusion policies deriving from suggested interventions. This stage can also include the compilation of the findings and insights into a comprehensive report of the interactions and interventions.

We already presented in Section 2 the concepts and measures that make up the Data Analysis and Visualisation and the Evaluation stages of the framework – we did not describe the reporting part of the last stage, since it is highly dependent on the application context of the single studies.

3. We deem it not pratical to define multi-step strategies since each intervention step can modify the social network under study, requiring a subsequent round of network reconstruction (following the IoT-based methodology presented in Section 3) to preserve its validity.

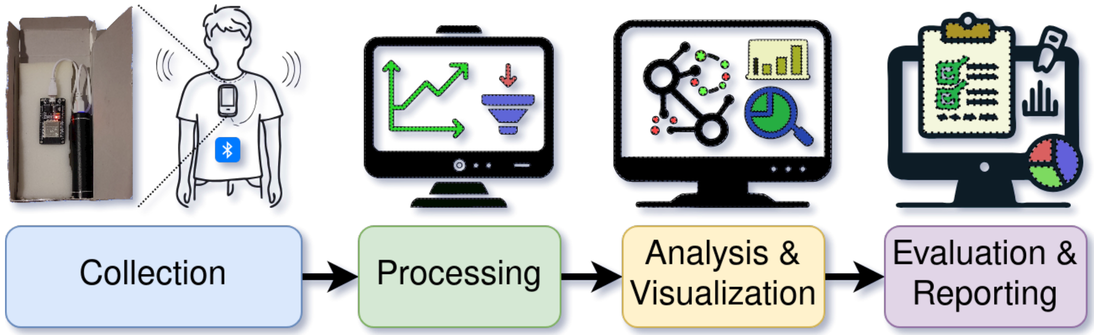


Fig. 3: Framework pipeline.

Hence, we dedicate the remainder of this section to the first two stages, i.e., that of data collection and processing.

3.2 Data Collection

The data collection stage uses IoT devices to gather social contact information for the later stages. We start by describing the prototyping of these devices and then move on to detail their usage in our framework.

3.2.1 Device Prototyping

Given the need for a cost-effective, child-friendly, and privacy-conscious method of data collection, we opted to develop a custom hardware prototype specifically designed for school environments. The decision was driven primarily by privacy concerns and the complexity involved with video-based methods, as well as the impracticality of devices requiring additional infrastructure such as base stations.

Since no suitable market options met our requirements, we built our own prototype using affordable, readily available IoT components. Specifically, we chose the ESP32 microcontroller for its optimal balance between performance, cost, and compact design, making it unobtrusive and comfortable for children to wear (see Fig. 3). Bluetooth Low Energy (BLE) technology was employed for its proven effectiveness in calculating proximity using the Received Signal Strength Indicator (RSSI), enabling accurate tracking of face-to-face interactions [55], [56], [57]. To achieve directional sensitivity and enhance detection accuracy, we designed and constructed a reflective shielding case. Given the ESP32's omnidirectional BLE antenna, this case blocks signals from behind and the sides, while increasing front signal reception within approximately 120 degrees, aligning with theoretical face-to-face interaction angles. Students wear the device on their chest, with the reflective front oriented outward.

The final prototype features a lightweight, cost-effective case made of cardboard, covered with aluminium shielding, foam padding for component protection, and a height-adjustable lanyard for comfortable wear. Internally, it integrates the ESP32 microcontroller, 4MB flash memory for efficient data storage, and a 1500mAh lithium battery sufficient for a two-hour session of continuous operation.

3.2.2 IoT-based Contact Detection

Once defined the hardware for tracking social contacts, we need to test its behaviour to discriminate between different contact configurations based on signal strengths. Our objective is to record the frequency and duration of social contact via Bluetooth radio waves. We use signal strength – a key metric in radio frequency engineering – to infer proximity, where stronger signals, measured in dBm, indicate closer distances. The devices operate in dual mode, broadcasting and scanning signals simultaneously, to continuously monitor interactions.

We set a sample rate of 1Hz for collecting all the signals within a device's range. We choose this sample rate for three main reasons: first, in a school setting, we do not expect significant sub-second movements between any two samples; second, the 1Hz sample rate strikes a balance between good data granularity and manageable data volume and energy consumption, providing the devices with over 2 hours of autonomy; third, as we detail in the next subsection, our definition of a valid contact requires a continuous interaction persisting for at least 5 seconds, making a 1Hz sampling rate sufficient to reliably capture such events even in the presence of occasional packet loss. To establish the optimal signal strength threshold for valid contacts, we conducted a series of static and dynamic tests in different configurations. These configurations, ranging from two subjects closely facing each other to more distant and angled arrangements, helped us fine-tune the device's sensitivity to various interaction dynamics.

Focusing on static tests, we recorded a variety of configurations, reported in Table 1, by positioning two testers, each wearing a front-facing device on their chest, at specific static distances from one another and recording the Bluetooth RSSI values over time. The aim of analysing this data, as summarised in Table 2, is twofold: to assess the devices' operational functionality and to establish an RSSI threshold indicative of legitimate face-to-face contact. In Table 2, we report the RSSI average and standard deviation recorded by the devices on a 60-second window wrt the considered configurations from Table 1, where the subjects stood at different distances apart – namely, 60, 80, 100, and 120cm, the latter being the farthest within the subject's personal space. We made a total of 11 distinct configurations, labelled A through M, each designed to test varying distances, angles, and orientations of the wearers' configurations. We report the specifics of each configuration in Table 1. Notably, in configurations where the devices were not aligned on the same horizontal plane (for example, configurations B, E, and F, among others), we introduced a vertical separation of

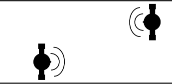
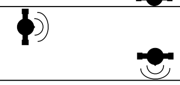
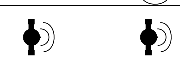
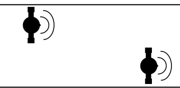
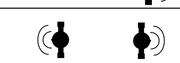
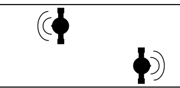
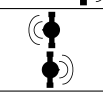
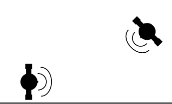
Conf.	Representation	Description	Validity
A		Face-to-face	✓
B		Face-to-face, 60cm shift	✓
C		Orthogonal, up	✗
D		Orthogonal, down	✗
E		Orthogonal, up, 60cm shift	✗
F		Orthogonal, down, 60cm shift	✗
G		Face-to-back	✗
H		Face-to-back, 60cm shift	✗
I		Back-to-back	✗
L		Back-to-back, 60cm shift	✗
M		Side-by-side	✗
N		Face-to-face, 45° angle, 60cm shift	✓

TABLE 1: Test configurations.

Conf.	Dist.	Avg	StDv
A	60cm	-47.15	5.08
A	80cm	-59.24	2.01
A	100cm	-57.36	14.34
A	120cm	-57.00	5.00
B	60cm	-66,69	1.48
B	80cm	-60.58	4.36
B	100cm	-71.10	7.62
C	100cm	-77.78	4.94
D	100cm	-75.26	2.44
D	120cm	-75.00	2.50
E	100cm	-72.12	6.39
E	60cm	-57.73	6.96
F	100cm	-71.63	3.02
G	60cm	-64.62	7.44
G	100cm	-68.41	1.98
H	100cm	-72.09	5.70
I	120cm	-71.00	7.00
L	60cm	-71.26	6.99
M	60cm	-67.68	8.90
M	80cm	-61.47	1.30
M	100cm	-62.71	2.30
N	120cm	-62.00	2.30

TABLE 2: Test configurations wrt different distances with RSSI average and standard deviation within a 60-second window.

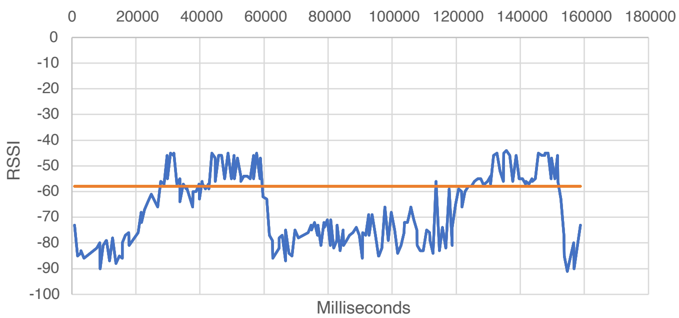
60 cm to further diversify our testing conditions. To clarify the setup, in configuration M, we refer to an arrangement where a person is oriented towards the right, while the other is directed at a 45-degree angle towards the bottom-left. The two are separated by a vertical distance of 60 cm. From the gathered data, we can assert that the use of the device within the shielded case is effective in detecting face-to-face contact. The RSSI values exhibit significant variance between valid configurations (where individuals are facing each other) and invalid ones (we mark the valid and invalid ones in Table 1 resp. with the symbols ✓ and ✗).

Besides static tests, we ran dynamic ones to evaluate the device's performance in changing configurations, where subjects are moving and interacting according to a scripted sequence of movements. These tests play a crucial role in our empirical approach to identifying the signal strength threshold for valid contacts, allowing us to fine-tune the signal strength threshold established during the static tests and increasing our confidence in the device's accuracy in real-world conditions.

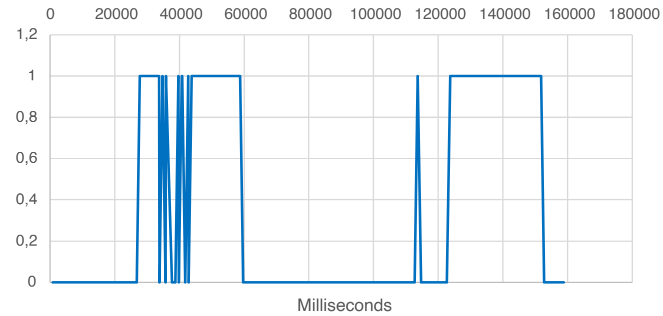
For the dynamic tests, we simulate real-life configurations by having two individuals move within a space, following a scripted action plan that dictates their movements over time. This script is carefully designed to include a range of interactions, from valid face-to-face contacts to configurations where no direct interaction occurs, within 120cm. The sequence begins with a 30-second preparation phase, mimicking configuration I (back-to-back positioning), followed by 30 seconds of valid interaction configurations (A and B), 60 seconds involving a mix of invalid configurations (C, D, E, F), and concluding with another 30-second phase of valid configurations.

Throughout these dynamic tests, we recorded timestamps and Bluetooth RSSI values to analyse the device's ability to track and detect interactions accurately. The analysis focuses particularly on the timeframes designated for valid configurations, aiming to validate or adjust the RSSI threshold determined from the static tests.

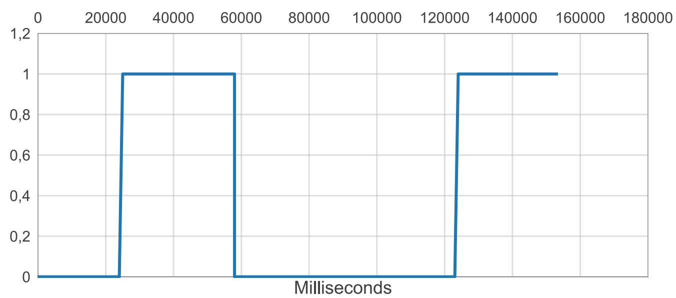
Fig. 4a illustrates the results of a sample of these recordings. In particular, we focus on two 30-second time frames where we tracked valid configurations. During these periods, we observe RSSI values occasionally dropping below the initial static configuration threshold of -57dBm, attributed to the unstable nature of Bluetooth signals and the irregularity of human movements. When analysing these valid contact timeframes, we adjust the RSSI threshold to -58dBm to enhance contact sensitivity, as indicated by the red line in Fig. 4a. This comprehensive approach, combining static and dynamic testing, enables us to refine the device's settings for optimal performance in detecting and recording social interactions. By analysing the data from these tests, we establish a more reliable and accurate signal strength threshold that reflects the complexities of real-world social dynamics. Once we terminated this battery of tests, we established the RSSI threshold for a valid contact at -58dBm. Note that, on average, this threshold excludes the invalid configurations from Table 1, although it might also exclude valid N ones, which tend to have a lower RSSI. We took this choice to ward off false positives while accepting a reasonable amount of false negatives (e.g., distant yet valid N configurations that we discard). To ensure robustness in real deployments, we complement this choice with a



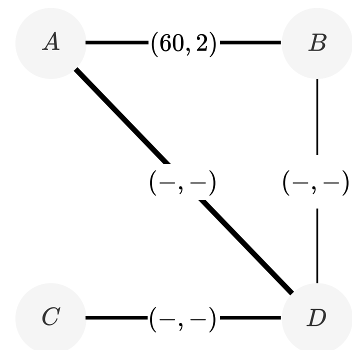
(a) Collected data.



(b) Threshold filtering.



(c) Noise removal.



(d) Network assembly.

Fig. 4: Example of data processing stages.

light on-site calibration: at the start of each deployment, we run a brief (2–3 min) check at known spacing ranges (considering the distances reported in Table 2) to verify device consistency and environmental conditions. In this way, we ensure periodic re-checks for long-running studies. In multi-week deployments, a weekly sanity check and, where necessary, a small additive bias correction can be applied to maintain threshold consistency across devices without altering the data pipeline. This procedure preserves comparability across sessions while guarding against device-to-device variability and environmental drift.

3.3 Data Processing

Once we have collected the data from the devices, we need to convert it into social contacts and social interactions. In the Data Processing stage, the collected raw data from the devices undergo a transformation, converting them into a structured format, suitable for the analysis found at the later stages. Concretely, we refine the data from the devices, which may contain signal errors, false positives, and hardware-dependent inconsistencies, into a reliable graph of face-to-face contacts and interactions.

The processing begins with a thorough cleaning of the raw data, where anomalies due to signal errors or hardware malfunctions are identified and rectified. The core of this stage is the definition/application of an algorithm designed to interpret the Bluetooth signal data, which follows the steps below:

1) *Signal Strength Analysis*: this phase filters out data points that fall below the established RSSI threshold, ensuring that only interactions within the defined proximity range are considered. Fig. 4b illustrates the application of this threshold check to our test data wrt the original collected data from Fig. 4a.
2) *Temporal Analysis*: we aggregate contacts that occur within a given time frame, thereby identifying continuous ones. This process is key in differentiating between meaningful contacts and spurious ones. For our purposes, we deem a contact valid if it persists for a minimum duration of 5 sec. This threshold is based on our framework's definition of *contact threshold*, described in Section 2.1, and the need to account for the instability of the RSSI of BLE signals. Empirical testing revealed that a 5-second threshold optimally balances the need to filter out spurious contacts while reliably capturing substantial face-to-face contacts. Fig. 4c illustrates the implementation of this temporal analysis, where, e.g., we aggregate the contacts within the ca. 22- and 59-second and the ca. 115- and 150-second marks from Fig. 4b. Note that, in Fig. 4c, the plot stops 20 seconds earlier (at ca. the 150-second mark) than in Fig. 4a and Fig. 4b. The reason for the "missing" values in the plot derives from the 20-second aggregation window used by the noise removal algorithm, i.e., to observe the "closure" of the interaction started at the ca. 115-second mark we would need the full 20-second window, which

would reach the 170-second mark.

3) *Interaction Validation*: we reconcile each potential interaction by checking whether a given couple of devices agree on having recorded their reciprocal RSSI. There are various ways to perform this task. We choose to follow the most inclusive policy, i.e., for each recorded interaction between any two devices, we take the maximum RSSI recorded. We implement this logic since we want to record all bidirectional interactions, even if they are recorded only by one device.
4) *Graph Construction*: we proceed to build the graph we described in Section 2 from the valid interactions – where each node represents a student and edges are the relationships between them, weighted by the number of their interactions (frequency) and the sum of their total time (duration). Fig. 4d shows a visualisation of one such graph, where we represent the recorded signal from Fig. 4a as device A receiving the signal of device B (we add the other nodes as an example and, thus, leave their weights unspecified).

By applying this algorithm, we transform the raw data into a graph that accurately represents the face-to-face interactions among students. As mentioned, this graph serves as the reference data for the subsequent stages of our framework, i.e., *Data Analysis and Visualisation* and *Evaluation and Reporting*, enabling us to analyse the social network of students and identify key patterns and relationships.

An alternative approach to interaction detection

Above, we define as part of the proposed framework an algorithm to detect contacts and interactions. Since this item is quite relevant to the validity of the dataset collected by the devices, we consider an alternative algorithm proposed by Barrat et al. [58] and use our case study (cf. Section 4) to compare its performance against the one we propose. Broadly, the main difference with our algorithm is that Barrat et al. considers as valid contacts those where there are at least five valid detections within a 20-second time frame. Given this stricter logic to contact-to-interaction conversion, we expect Barrat et al. to drop short-term interactions and record mostly longer ones.

4 Case Study

We start the validation process of our framework by applying it to a case study of primary school students. As per diligence, we describe the participant selection process and the methodology used to conduct the experimentation. After the in-field data collection, achieved by deploying one prototype for each student, we process the data through the data pipeline presented in Section 3. In particular, since we record the raw data of the signals received by the devices, we evaluate our algorithm for interaction detection wrt an alternative one, i.e., the one defined by Barrat et al. – since the devices preserve the detected device ID, timestamp, and signal strength, we can apply both algorithms after the gathering of the data during the experiment. Hence, we conduct the subsequent steps of our social inclusion routine and proposal for intervention on both networks, drawing conclusions both at the level of the case study and, at the meta-level, on the proposed framework – commenting on the feasibility of the deployment and the overall process of data handling and suggestions on interventions.

Participants

The study subjects are 16 students (7–8 years old), 8 male and 8 female, from a third-year elementary class at a school in Bologna county, Emilia-Romagna region, Italy. In adherence to the ethical guidelines of the affiliation of the authors, the University ethics committee reviewed and approved the study project, which the hosting school also approved, obtaining signed consent from the parents of all the subjects involved in the study – all involved parties obtained and signed documentation about the aims of the study and provided informed consent.

Data Collection

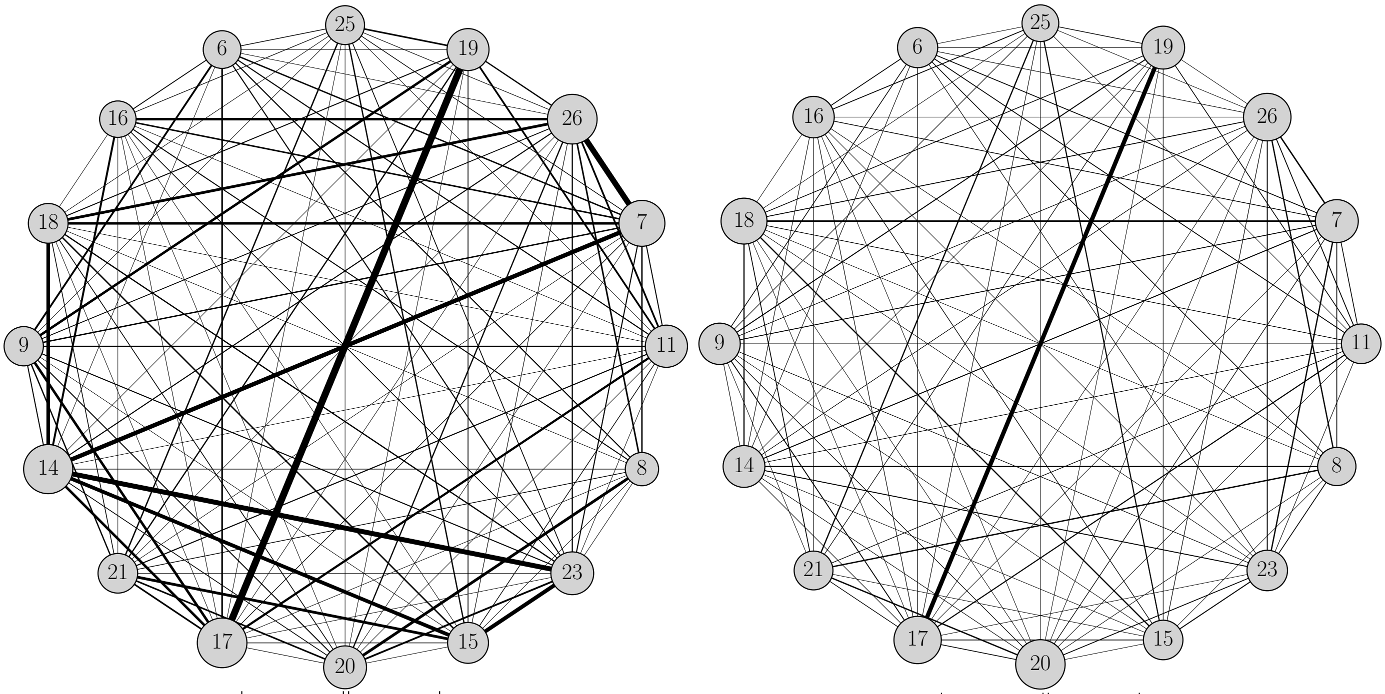
Following standard procedure, we instructed one of the teachers (who the students are familiar with) to explain to the students the purpose of the study (gather information on the interactions they have) and the rules of the study.

The experiment took place during a two-hour recess. The teachers gave the students a room where they were free to roam around and do what they wanted – the researcher had previously prepared some desks in the room with materials such as puzzles, cards, sheets, and pencils, which are normally available during recess. Before beginning the experiment, the teacher indicated that running, jumping, and exchanging the devices was forbidden during the experiment. At the beginning of the experiment, one student at a time extracted a device from an urn (to ensure anonymity), the researcher helped them wear it correctly (checking that the device was firmly held by the straps and adjusting the latter to have all devices at around the same height for all students), and switched the device on. Once a student had a switched-on device, they could freely engage in the activities they liked. Since the deployment of the devices followed a sequential process, the researcher recorded the time when they switched on the last device. This datum indicates the starting time of the experiment, i.e., when we start considering the recorded interactions among the subjects.

Both the teacher and the researcher stayed in a corner of the room of the experiment to ensure the students did not violate the rules – yet without interacting with them or influencing their behaviour unless they violate any rule, which did not happen. After two hours, the researcher recorded the time of the end of the experiment – which marks the time since when we discard further interactions recorded by the devices – and proceeded to gather and switch off the devices.

Data Processing and Results

After the data collection phase, we proceeded to extract from each device the log file containing the information on the detected device IDs, timestamps, and RSSI. Once we gathered all the logs from the devices, we followed the steps of the pipeline (cf. Section 3), starting from the application of our algorithm to normalise the recorded contacts and aggregate them into interactions. In this phase, we generate a second dataset, obtained by applying the algorithm by Barrat et al., for comparison.



19-17	184.5	26-7	149.83
23-14	131.14	14-7	113.8
23-15	104.0	15-14	99.0
18-14	95.11	20-8	74.66
21-15	73.25	26-18	72.87

19-17	114.12	18-7	39.63
26-7	37.15	21-20	35.14
21-8	35.0	23-7	34.4
26-8	33.66	25-21	33.2
18-15	33.0	18-14	31.9

Fig. 5: Left, the network's graph generated from our algorithm, right, the one using Barrat et al.. We report at the bottom of each graph the 10 strongest relationships.

We report the obtained data at https://zenodo.org/records/12948919, where we also provide the Python source code that implements our social inclusion algorithm, used to analyse this case study. For clarity, node labels in Figure 5 and Table 3 reflect device identifiers rather than a contiguous participant index. Devices were pre-provisioned from a larger pool (IDs up to 26), but only 16 were actively assigned to participants during the session. We retain the original device IDs to preserve data provenance across acquisition, processing, and analysis.

In Fig. 5, on the left, we show the network obtained via our algorithm. In the remainder, we call this network G. On the right of Fig. 5, we show the one generated by Barrat et al.; from now on, called G'. Both graphs follow the layout of Fig. 2, where, for compactness, we already apply the measures of our framework (cf. Section 2), sizing resp. the nodes and the edges proportionally to their ι and duration/frequency ratio. To make a more detailed comparison between the networks generated via the two algorithms, we report the 10 highest-valued edges at the bottom of each graph in Fig. 5.

The comparison shows that our algorithm seems more fine-grained than the one by Barrat et al., although they both tend to detect the interactions among the subjects similarly. Specifically, we notice how the only "relevant" strong interaction in G' (in terms of the ratio between time and frequency) is the one between 19 and 17, while the other edges are in the same bracket. Contrarily, our algorithm gives more refined interaction values. Specifically, it agrees with Barrat et al. in assigning the strongest value to the relationship between 19 and 17, but other strong ones emerge, like 26-7 and 23-14 – in general, our algorithm suggests that 14 is an important node that has many strong relationships with other nodes in the network; an observation that is harder to make in G'. The visualisation matches the data reported at the bottom of Fig. 5, where, besides the 19-17 edge, all strongest relationships in G' have nearly the same value. Contrarily, the strength of the relationships determined by our algorithm seem to be more granularly detected. In general, we notice that the main difference between our algorithm and the one by Barrat et al. is that the former detects more connections (generating a lower cumulative frequency), resulting in a longer cumulative duration of interactions. We expected this outcome, given the timeout system we introduced in our algorithm to account for interactions that extend over long periods but present brief interruptions (seconds) in the logs of recorded signals.

Data Analysis

Moving to the analysis of the data, G has 104 relationships and 724 interaction events. The length of paths therein has an average of 1.13, with a diameter of 2. The students have an average degree centrality of 13 and interact with classmates 6.96 times for a cumulative interaction time average of 256.54 seconds. The interaction frequency ranges from 2 to 40, while the duration ranges from 2 to 1755 seconds. In contrast, G' includes 96 relationships and 889 interaction events and has an average path length of 1.2 and a diameter

n. id	ι	n. id	ι	n. id	ι	n. id	ι
8	0.0433	9	0.0603	19	0.0688	7	0.0824
16	0.0519	21	0.0614	11	0.0696	14	0.0944
6	0.0557	18	0.0617	23	0.0716	26	0.0962
25	0.0591	15	0.0653	20	0.0718	17	0.0962
n. id	ι	n. id	ι	n. id	ι	n. id	ι
25	0.0530	11	0.0614	16	0.0672	18	0.0826
8	0.0575	6	0.0632	14	0.0685	17	0.0878
21	0.0590	23	0.0643	19	0.0720	26	0.0884
15	0.0612	9	0.0669	7	0.0723	20	0.0958

TABLE 3: The ι values of all nodes, sorted from lowest (rightmost column) to highest (leftmost column). Top, values for the network from our algorithm, bottom, for Barrat et al.

of 2. Students have an average degree centrality of 12 and interact on average with classmates 9.26 times for 198.44 seconds per session. The frequency of interactions ranges from 2 to 61, while the duration ranges from 2 to 1149 seconds. Summarising the results from the case study, we notice that Barrat et al. does to convert many "weaker" contacts into interactions because, once the signal between two devices is lost, it does not consider close future contacts to establish an interaction. Hence, that algorithm tends to recognise as interactions stable and prologued contacts, like the ones happening between 19 and 17, while the "weaker" interactions gather around the same values. Contrarily, since our algorithm compounds contacts within the given contact threshold, it can recognise shorter-term interactions and, consequently, it can weight interactions in finer terms than Barrat et al..

We apply our measures routine to find the most isolated node and propose a policy to increase their integration into the network. We report in Table 3 the resulting ι values for the nodes of G (top) and G' (bottom).

Notably, the ι values for the nodes in the network identify the nodes in similar ways. Indeed, while the most isolated node is 8 in G and 25 in G', 8 is the second most isolated node in G' and 25 is the fourth most isolated node in G. Broadening our scope, we find an overlapping set of nodes at the lower half of the ι spectrum, except for 16 and 18 in G, which are replaced by 11 and 23 in G' – and a similar observation holds for the higher half.

Proceeding with the application of our intervention procedure, we identify 8 as the most isolated in G ($\iota = 0.043$) and calculate all possible integration scenarios. Among these, the best-scoring one is connecting 8 to 19 (a well-connected node, but the lowest of the most integrated ones according to Table 3) thanks to node 26, which grants a frequency of 3 and a duration of 124 seconds. The intervention raises 8's ι to 0.054, for a 26.42% increase. From the point of view of the "stress" exerted on the network by the intervention, the assortativity coefficient by degree of G is -0.088 (indicating a slight preference of the nodes to interact with nodes that are more "connected" than they are), which is slightly lowered to -0.073 (ca. 17% less) by the intervention.

For comparison, the most isolated node in G' is 25, with a ι of 0.053. The best-scoring intervention scenario for G' and 25 is to connect the latter to 8 (interestingly, the second most isolated node in G', cf. Table 3) through 21, which affords a frequency of 1 and a duration 21 seconds. That intervention raises 25's ι to 0.064 (a ca. 22% increase), impacting minimally the natural tendency of the nodes in the network to establish relationships, G''s assortativity coefficient by degree is -0.068 (showing the tendency of nodes to establish connections with more "central" ones seen in G) which is slightly increased by the intervention to -0.070 (ca. 3%). For completeness, we visualise both G and G' after the respective interventions in Fig. 6, marking the new relationships as dashed red lines.

The case study shows how the IoT devices can be a powerful tool for identifying the social network within a classroom, as it also has the potential to detect students who are central to the network as well as those who are at risk of social isolation. Social isolation could be detrimental to these students, since the feelings of belonging and peer acceptance are crucial to students' emotional well-being as well as for their school performance [59]. The findings from this case study highlight the potential of IoT devices in educational research and policy design. In fact, it is crucial to emphasise that identifying the relationships within the classroom is the first step, and translating the findings into specific interventions to promote social inclusion is also an aim of our study. More precisely, the challenge for these interventions lies in attempting to change the network dynamics within the classroom and to foster meaningful social relationships (see Section 6 for a detailed description of these interventions). Proposed interventions should be able to encourage collaborative activities and allow sufficient time for social bonding to take place in order to promote social inclusion.

Abstracting from the case study and looking at the application of our proposed framework, we deem the successful completion of all the phases of the experiment initial proof that our proposal is valid and feasible for the context considered in the premises and this experiment (e.g., students in primary school). As mentioned, the framework does not propose a policy for the application of the suggested intervention, which is left to the interpretation of a domain expert, who must consider the context in which the intervention is proposed for its reification, e.g., what activity to do, within what time window to apply the policy and allow the change to settle, which also determines when it is possible to repeat the experiment to detect the modifications induced by the intervention.

5 Related Work

To the best of our knowledge, this is the first work that marries SNA and the usage of IoT devices for the detection and improvement of social integration. Hence, while we cannot compare with work from the literature directly, we position our work wrt existing contributions on IoT-enabled Social Network Analysis, which focus on the technological advancements and methodologies for collecting and analysing interaction data, and Social Network Analysis for studying social integration, which studies how one can use these technologies and methods to foster social inclusion.

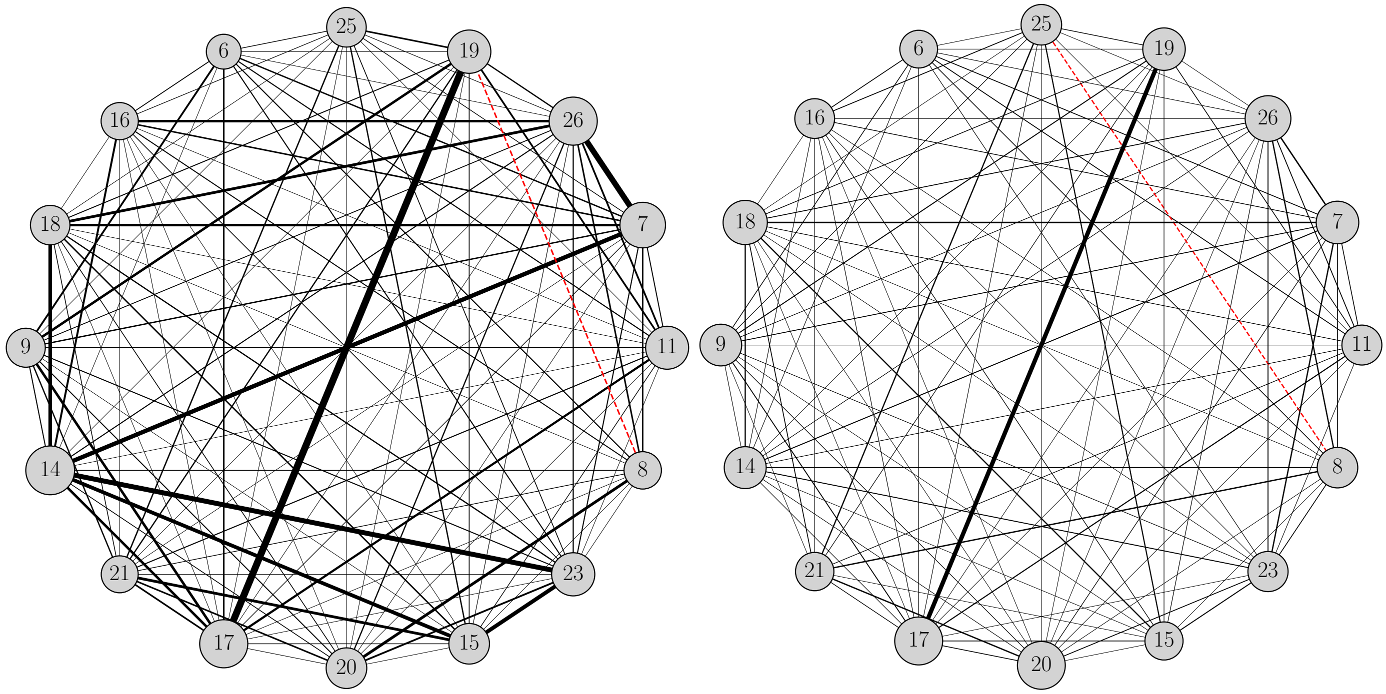


Fig. 6: Proposed integration (dashed red). On the left for the network obtained with our algorithm and on the right for the network obtained from Barrat et al.

5.1 IoT-Enabled Social Network Analysis

The integration of IoT into SNA has significantly advanced the ability to gather detailed and real-time data on social interactions. Early work in this area, such as [60], used RFID devices to track person-to-person interactions, centralising data collection through routers. Although this approach demonstrated the feasibility of IoT for interaction detection, its infrastructure requirements posed scalability challenges. In particular, we notice that, wrt the other techniques which require the installation of stable infrastructural nodes, like routers and fixed beacons, the solution we propose is decentralised and requires no intervention on the environment of the experiment. Building on this foundation, the SocioPatterns project [21] has been pivotal in demonstrating the effectiveness of wearable sensors in capturing interaction patterns across various settings, including schools and public health scenarios. These studies underscore the robustness of IoT devices in practical applications, despite their higher costs and complex installations compared to more recent low-cost BLE-based approaches.

Isella et al. [61] extended this analysis to large-scale real-world scenarios, comparing face-to-face interactions at a scientific conference and a museum exhibition. Their findings illustrated the dynamic nature of human interactions and their impact on epidemic-spreading models, emphasising the need for detailed interaction data to understand transmission paths. Similarly, Stehlé et al. [62] explored the impact of social relationships on student academic performance and social well-being, highlighting the critical role of peer interactions in shaping educational outcomes. Their research provides a strong rationale for interventions aimed at enhancing social inclusion. Exploring the application of wearable sensors, Colosi et al. [63] investigated the spread of infectious diseases through social networks by collecting close-proximity interaction data. Their analysis modelled the potential spread of infections, highlighting the critical role of accurate interaction data in both social and epidemiological studies. While we use precise interaction tracking to enhance social inclusion, the techniques Colosi et al. use align with ours.

To validate the approach, Mastrandrea et al. [64] compared different data collection methods, including wearable sensors and contact diaries, in a French high school. This study revealed that short contacts are often underreported in diaries, while long contacts have a high reporting probability. Despite biases in self-reported data, the overall structure of the contact network was accurately captured, showcasing the complementary nature of sensor-based data and traditional survey methods. Interestingly, wearable proximity sensors have shown significant potential in constrained-resource settings. For instance, Kiti et al. [65] used these sensors in rural Kenya to study face-to-face interactions within households, revealing stable interaction patterns over several days. Their study also emphasised the critical role of community engagement and privacy considerations in ensuring the acceptability of such devices in these environments. Hasan [66] provides a recent systematic review of IoT-based smart education systems, highlighting their potential to improve accessibility and social inclusion in educational environments. The study categorizes 24 recent contributions and outlines both technical advancements and socio-educational impacts. Further, wearable devices have been validated in recent field experiments to collect fine-grained contact data for SNA applications [67], supporting their reliability in dynamic school settings.

A recent comparative analysis [68] evaluated different proximity sensing technologies for social contact detection in schools. It found BLE-based solutions to be more effective in supporting both educational and epidemiological studies, reinforcing the technical foundation of our work. Lastly, Thompson et al. [69] explored how wearable sensors can identify at-risk students based on social connectivity levels, providing experimental validation for early interventions – a direction strongly aligned with our framework.

Our framework integrates affordable IoT devices with advanced data processing algorithms, offering a novel approach to measuring and enhancing social inclusion in primary schools. This method addresses the specific challenges and opportunities associated with studying young children's interactions, providing more granular and objective data compared to traditional observational methods.

5.2 Social Network Analysis for Social Inclusion

SNA has been instrumental in understanding and enhancing social dynamics, particularly in educational settings.

Berkman and Glass [70] provide a landmark summary regarding work on the development of the field of sociology and network analysis involved in the study of social integration and support. In particular, Berkman and Glass underline how SNA can be instrumental in detecting the influence on integration of both the macro level, i.e., socio-structural conditions like norms and values but also discrimination, economic status, and social cohesion, as well as the micro level, i.e., psychological mechanisms like peer pressure and social roles.

Focussing on recent applications of SNA related to social inclusion, Abbott et al. [71] showed the feasibility of using SNA to model the relationships of older adults in an assisted living neighbourhood, showcasing network visualisation (sociograms) to illustrate the level of social integration of the studied subjects and network measures to evaluate a person's network cohesion. Recent studies have explored participatory SNA approaches to uncover the emergence of social dynamics in rural development contexts, emphasising the role of community-driven data collection [72]. While set outside of education, the method aligns with our goal of empowering local stakeholders with contextualised interaction data. Similarly, Garrote et al. [73] used SNA to examine peer learning strategies among students, revealing how the analysis of collaborative patterns can inform inclusive pedagogical design. Karimi and Matous [74] applied SNA to map students' social activities and indicate how SNA's measures can inform university policies aimed at increasing social inclusion. Pearson et al. [75] used SNA to understand how students from diverse backgrounds integrate into introductory engineering courses. Finally, Singh et al. [76] recently provided a broad overview of SNA in education, classifying its applications across peer learning, collaboration, and inclusion. Their findings align with our goal of using interaction-based data to inform inclusive practices.

Our study contributes to the body of work on social inclusion by proposing a novel measure of social inclusion (ι) and an algorithm that proposes integration interventions balanced wrt the tendency of the subjects in the network to constitute relationships with one another. The integration of SNA techniques with IoT-based contact tracing in the setting of education addresses the distinctive challenges of studying young children's interactions, balancing cost, scalability, and accuracy, and making it a viable option for various educational contexts.

An additional lens through which we can consider work related to ours is that of link prediction [77], i.e., finding the next most probable links that the members of a network can establish, according to the network's topology and the members' properties, and link recommendation [78], i.e., estimating the value of potential links and recommend the establishment of the ones with the highest scores. We deem our contribution mostly aligned with this second branch of research, where the value we attribute to the new links is their impact on improving the level of social inclusion of the most isolated nodes. Considering this area of research, we foresee future work concerning the refinement of the current logic of link establishment suggestion, e.g., considering additional aspects such as the clustering of nodes [79] and attribute-based homophily [80]. More recently, Bashardoust et al. [81] introduced measures that represent different traits of a node's ability to share information with other nodes, considering the problem of improving equity by making interventions to improve access for the least advantaged nodes. The main difference between our work and the one by Bashardoust et al. is that the latter considers a general notion of a node's (dis)advantage position based on its ability to broadcast, influence, and control the network, while we specifically consider the problem of social inclusion and define a methodology to study and propose suggestions based on context-specific network traits (like IoT-based network structure detection, the definition of network measures based on duration and frequency of the interactions, etc.). Interestingly, the proposal by Bashardoust et al. can inspire future evolutions of our contribution, e.g., by integrating and grounding within the context of social inclusion the measures they propose.

6 Discussion and Future Work

We presented a framework to quantitatively measure social inclusion and suggest actions to increase the integration of more isolated agents.

In the theoretical part (Section 2), we drew upon social theories and human behaviour to introduce methods for measuring social relations and network analysis measures for assessing the level of inclusion and solution quality. Given a network representation of the studied agents, we build a network model and define network measures to indicate the less integrated agents and propose sustainable interventions to increase their inclusion.

To lower the barriers to entry for researchers in the field, we base our framework on over-the-counter IoT devices for detecting the frequency and duration of social interactions. In the practical section (Section 3), we illustrated the development of a device prototype for data collection to apply this framework in a real-world scenario. Then, we presented a pipeline and algorithms useful to manage the data collected from the devices and build the network where one can apply the network measures and procedures for social integration.

To validate our proposal, we conducted an in-field case study on a class of primary school students, practically demonstrating the feasibility of the proposed approach, in particular in a context where investigating social interactions might be more difficult via traditional methods like questionnaires or interviews.

Limitations and Future Work

From a validation standpoint, our study offers an initial confirmation of the framework's potential. Yet, we acknowledge that further research is necessary to establish its robustness both within the same context and across different ones. In particular, we recognise that the current validation is limited due to its focus on one sample and that additional studies – including replication ones, with larger and diverse samples – are essential to bolster the statistical evidence and generalisability of our approach.

Interestingly, we note that extending the usage of our framework to different environments, such as workplace settings, would require only minor adjustments, primarily in parameter configuration and the design of the experimental protocol. This factor confirms the framework's core structure inherent adaptability. These future studies are essential to address the need for contextual fine-tuning and to systematically examine the potential criticalities related to transferability and applicability of the approach.

Beyond the items above, we note two important points that remain to be addressed in future work: *a*) the application to case studies that apply the integration policies implementing the suggestion of the framework and *b*) the conduction of comparative studies that evaluate the accuracy and reliability of the detected contacts by the devices and the validity of the measures of integration.

To address the first point, one can define and run case studies that are extensions of the one conducted in Section 4, which add, on top of the latter, two new phases, for a total of three. The first phase, called *ex-ante*, records the current configuration of the network of agents under study and applies our integration algorithm to identify the less integrated nodes and obtain a suggested target intervention to increase social inclusion. The second phase, dubbed *integration policies*, regards the application of social policies designed to achieve the configuration suggested by the framework in the previous phase. This phase is orthogonal to the application of the framework and highly dependent on the context. We discuss this matter in the next paragraph, on policy suggestions. The third phase, called *ex-post*, repeats the experiment of the *ex-ante* to verify whether and how much the *integration policies* succeeded in changing the topology of the network towards the goal configuration of the *ex-ante*. In connection to these repeated observational phases, an important assumption underlying our framework is that individuals' natural interaction tendencies remain relatively stable over the short-term period considered for interventions. However, social dynamics can change significantly over longer periods, potentially affecting the long-term effectiveness of our suggested interventions. To address this point, we recommend periodically repeating observational studies to assess network stability and detect changes in interaction patterns over time. Understanding how and why these patterns evolve would allow practitioners to adapt intervention strategies accordingly, thus maintaining the effectiveness of the framework despite shifts in social dynamics. Beyond the short-term stability assumption, we also acknowledge limitations in terms of *external validity*. The framework has so far been validated in a single primary school setting, where social dynamics are relatively homogeneous and structured. In other contexts – such as workplaces, mixed-age groups, or culturally diverse classrooms – interaction patterns may evolve more rapidly or follow different norms, which could reduce the direct transferability of our results. Therefore, future work should examine the stability of detected interaction patterns across heterogeneous environments and over longer time horizons, to ensure that the proposed measures and interventions are robust and generalisable.

In response to the second point, one can run a comparative study adopting a two-level approach. The first level validates contact detection against observational data. Hence, this level verifies how accurately the devices detect social contact between agents by comparing the latter against observational data, like video recordings of the agents during the experiment. Given these recordings, a set of researchers can document the time of each agent's social contacts and assemble a comparable dataset to the one recorded by the devices. Although laborious and resource-demanding (multi-angle camera recordings, time dedicated to transcribing the observed contacts and reconciling possible discrepancies), this data would provide solid evidence of the validity and accuracy of the social contact data recorded by the devices. The second level uses surveys to gather data regarding the social relationships among the studied agents. In this way, one can build a quantitative model of the integration of each agent – e.g., using a Likert scale to ask each respondent how integrated they feel and, complementarily, how the others "score" the latter's integration – and compare the obtained data against the integration score provided by our framework. Notably, besides the comparison, one can integrate the data obtained through our framework and the one from the survey to gain a more integral account of the integration levels of the agents.

Looking at the evolution of the framework, one could further refine the sensitivity and coverage of the measure of social inclusion, using alternative measures to the one we defined and/or introducing new ones such as clustering algorithms. Moreover, as mentioned in Section 2.2.2, we are interested in developing multistep intervention techniques that could improve the performance of the framework and solve edge cases. Another interesting direction is using other attributes than the degree of nodes to "balance" the impact of the suggestion, through assortativity. Indeed, ascriptive factors like ethnicity, socio-economic background and gender can substantially influence how people build relationships with one another. We plan to conduct future studies where we use our IoT devices to study the process of tie formation focusing on the mechanisms identified in the literature like homophily, reciprocity, and transitivity [82]. Integrating these aspects in the intervention suggestion algorithm could help design more sustainable policies that take into account how the agents interact with each other depending on these attributes.

A valuable extension of our framework involves aggregating individual-level interaction data to analyse higher-level group dynamics. By employing techniques such as community detection or clustering algorithms, our framework could identify distinct subgroups or clusters based on interaction patterns [83]. These clusters could then serve as focal points for targeted interventions at a collective rather than individual level, potentially enhancing the scalability and effectiveness of inclusion efforts across larger populations or more complex environments.

Another important consideration for future work involves deepening our approach to ethical concerns, particularly regarding data security and user authentication. Given that our framework involves minors and operates within educational contexts, ensuring robust protection mechanisms against unauthorised data access is critical. Future iterations should explicitly implement advanced security practices such as strong device authentication, encrypted data transmissions, and rigorous access control protocols [84].
Finally, an interesting future endeavour is exploring the industrialisation of the prototype to reduce cost and dimensions, possibly extending the current functionalities of the devices, e.g., incorporating a camera and facial/emotion recognition technology to trace interaction quality – where in-device facial/emotion recognition is fundamental to preserving the privacy of the agents while providing valuable, anonymised data to researchers. Addressing these aspects will further enhance the ethical acceptability and practical deployment of our IoT-based social inclusion framework.

In conclusion, we consider the results presented in this article to also advance the field of social inclusion in at least two significant ways. First, the framework, as stated above, introduces a quantitative method to measure social inclusion by integrating IoT with social network analysis. This represents a valuable methodological alternative to traditional survey-based approaches, enabling the collection of more granular and more "objective" data on network structures and interaction patterns. Future research should explore how this approach can be combined with self-reported data to capture both the structural and subjective dimensions of social inclusion. Second, the proposed framework offers a replicable and adaptable tool set that can support empirical evaluation of inclusion-focused policies, particularly in educational settings. By generating precise, time-stamped relational data, the framework enables researchers and practitioners to assess whether specific interventions have measurable effects on the integration of isolated individuals. Future studies could build on our contribution by integrating administrative, behavioural, or psychological indicators to evaluate the broader impact of such interventions and to test the robustness of the framework in diverse institutional contexts.

Policy suggestions

Our main focus in this article is to identify pupils who may be isolated from their peers in the classroom in a real situation. At this point, we might ask: "what are possible interventions to reduce social isolation?". There are at least three types of interventions that could be useful in this setting: inclusive classroom practices, extracurricular activities, and parental involvement. As highlighted in Section 4, the role of positive peer relationships is pivotal to promoting school achievement [85] and well-being [86]. Previous research demonstrates that modifying the network structure is feasible [87] and suggests that designing and implementing interventions aimed at changing the network structure in a classroom could have positive effects on school achievement [88].

Inclusive classroom practices can play a crucial role in reducing social isolation by fostering an environment where all pupils feel valued and supported. This type of intervention is also easier to organise as it can be embedded in existing activities within schools. Collaborative learning is one of the cornerstones of this approach: teachers divide students into small groups with different projects and activities that encourage interaction and mutual support [89]. In this way, socially isolated pupils can be encouraged to interact and build relationships with their peers. In a detailed literature review, Johnson and Johnson [90] reported that cooperative learning not only promotes academic achievement but also improves social interactions and relationships among students. Another strategy that can be implemented to reduce social isolation is the use of circle time. Pupils sitting in a circle are encouraged to discuss different topics and express their feelings. This practice, based on the theory of belonging [91], should promote the creation of a sense of community within the classroom, contributing to the well-being of students and reducing feelings of isolation.

Extracurricular activities can also be a useful tool to reduce social exclusion, as these activities have been shown to allow students to make friends and reduce social isolation outside the classroom [92]. These interventions are mainly implemented in Anglo-Saxon countries where school-organised extracurricular activities are more developed. The idea is that schools should encourage a range of activities to cater to the diverse interests of pupils and to facilitate links between pupils with similar interests.

The last kind of intervention relates to the involvement of parents. One can implement these policies by organising seminars to illustrate how parents can support their children's social development [93] or by organising events that bring together students, teachers, and parents to provide an informal setting where students can interact outside the classroom. In summary, inclusive classroom practices, extracurricular activities, and parental and community involvement can be effective interventions that schools can implement to reduce social isolation among students. Although these interventions are based on group activities they can easily be adapted to build connections between the more isolated students with students that are more central in the network, as pointed out in Section 2.2.2. More precisely, the size of the groups and the way they are formed can be manipulated to encourage contact between certain types of students.

Broader applications

While our main focus was on primary school pupils, the proposed framework can be easily extended to other contexts where social inclusion is critical. For instance, in workplace settings, it could be used to monitor interaction patterns within teams or departments, helping identify peripheral employees and assess the degree of integration among colleagues. This information could inform human resource practices and complement existing team-building strategies aimed at fostering organisational cohesion. Similarly, the framework could be valuable in elderly care settings, helping staff to identify residents at greater risk of social isolation and guiding the development of targeted interventions. In such contexts, proximity-based sensors are a particularly suitable approach, as residents may have cognitive limitations or experience stigma that hinders their willingness or ability to report feelings of loneliness. These examples highlight the broad applicability of our approach and suggest future research opportunities in various institutional settings. While these settings may present unique challenges, they could all benefit from the application of our framework on social inclusion.

References

[1] K. D. Haynes-Mendez, S. A. Nolan, L. N. Littleford, and L. M. Woolf, "Diversity, equity, inclusion, and internationalization: Past, present, and future of stp," *Teaching of Psychology*, vol. 0, no. 0, p. 00986283221126424, 0. [Online]. Available: https://doi.org/10.1177/00986283221126424

[2] P. Jurado de Los Santos, A.-J. Moreno-Guerrero, J.-A. Marín-Marín, and R. Soler Costa, "The term equity in education: A literature review with scientific mapping in web of science," *International Journal of Environmental Research and Public Health*, vol. 17, no. 10, p. 3526, 2020.

[3] K.-T. Lindner, S. Hassani, S. Schwab, C. Gerdenitsch, S. Kopp-Sixt, and A. Holzinger, "Promoting factors of social inclusion of students with special educational needs: Perspectives of parents, teachers, and students," *Frontiers in Education*, vol. 7, 2022. [Online]. Available: https://www.frontiersin.org/articles/10.3389/feduc.2022.773230

[4] E. Wesselmann, J. Wirth, and M. Bernstein, "Expectations of social inclusion and exclusion," *Frontiers in Psychology*, vol. 8, 2017. [Online]. Available: https://www.frontiersin.org/articles/10.3389/fpsyg.2017.00112

[5] E. D. Wesselmann, J. H. Wirth, and M. J. Bernstein, "Expectations of social inclusion and exclusion," *Frontiers in Psychology*, vol. 8, pp. 112–112, 2017.

[6] J. Blicharski, "Social inclusion and higher education." *Widening Participation & Lifelong Learning*, vol. 14, no. 2, pp. 79–81, August 2012.

[7] E. Commission, "Proposal for a joint employment report from the commission and the council," Nov 2022. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A52022DC0783

[8] ——, "New skills / social inclusion," Feb 2019. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=PI_EESC%3AEESC-2019-01610-AS&qid=1669382480149

[9] ——, "Council conclusions on enhancing the social inclusion of young people not in employment, education or training," Mar 2014. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52014XG0201%2802%29&qid=1669382480149

[10] ——, "Joint report on social protection and social inclusion," Jan 2008. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52008SC0091&qid=1669382480149

[11] ——, "Joint report on social protection and social inclusion," Mar 2007. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52007SC0329&qid=1669382480149

[12] ——, "Communication from the commission to the council, the european parliament, the european economic and social committee and the committee of the regions - joint report on social protection and social inclusion," Feb 2006. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52006DC0062&qid=1669382480149

[13] ——, "European parliament resolution on social inclusion in the new member states," Nov 2004. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52005IP0244&qid=1669382480149

[14] S. A. European Commission, Directorate-General for Employment and Inclusion, "Resolution of the council and of the representatives of the governments of the member states," Dec 2000. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A42000X1228&qid=1669382480149

[15] M. Koster, H. Nakken, S. J. Pijl, and E. Van Houten, "Being part of the peer group: A literature study focusing on the social dimension of inclusion in education," *International Journal of Inclusive Education*, vol. 13, no. 2, pp. 117–140, 2009.

[16] 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.

[17] T. J. Williamson, K. S. Thomas, N. I. Eisenberger, and A. L. Stanton, "Effects of social exclusion on cardiovascular and affective reactivity to a socially evaluative stressor," *International Journal of Behavioral Medicine*, vol. 25, pp. 410–420, 2018.

[18] K. D. Haynes-Mendez, S. A. Nolan, L. N. Littleford, and L. M. Woolf, "Effects of social exclusion on cognitive processes: Anticipated aloneness reduces intelligent thought," *Teaching of Psychology*, vol. 0, no. 0, p. 00986283221126424, 0. [Online]. Available: https://doi.org/10.1177/00986283221126424

[19] R. Jenny, "A-z of inclusion in early childhood / mary dickins," Maidenhead, 2014.

[20] S. Qiu, P. An, K. Kang, J. Hu, T. Han, and M. Rauterberg, "A review of data gathering methods for evaluating socially assistive systems," *Sensors*, vol. 22, no. 1, p. 82, 2021.

[21] A. Barrat, C. Cattuto, V. Colizza, J.-F. Pinton, W. V. d. Broeck, and A. Vespignani, "High resolution dynamical mapping of social interactions with active rfid," *arXiv preprint arXiv:0811.4170*, 2008.

[22] E. Avramidis, V. Strogilos, K. Aroni, and C. T. Kantaraki, "Using sociometric techniques to assess the social impacts of inclusion: Some methodological considerations," *Educational Research Review*, vol. 20, pp. 68–80, 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1747938X1630063X

[23] T. Coombs, A. Nicholas, and J. Pirkis, "A review of social inclusion measures," *Australian & New Zealand Journal of Psychiatry*, vol. 47, no. 10, pp. 906–919, 2013, pMID: 23737598. [Online]. Available: https://doi.org/10.1177/0004867413491161

[24] C. Mamas, "Employing social network analysis to examine the social participation of students identified as having special educational needs and disabilities." *International Journal of Disability, Development, and Education*, vol. 67, no. 4, pp. 393–409, July 2020.

[25] J. L. Moreno, *Sociometry, Experimental Method and the Science of Society. An Approach to a New Political Orientation*. Ambler (PA): Beacon House, 1951.

[26] R. K. Merton, *Social Theory and Social Structure*. New York: The Free Press, 1968.

[27] P. Sztompka, *Socjologia*. Kraków: Znak, 2002.

[28] B. S. Turner, *The Cambridge Dictionary of Sociology*. Cambridge: Cambridge University Press, 2006.

[29] A. D. of Psychology. social relationship. [Online]. Available: https://dictionary.apa.org/social-relationship

[30] J. Cooper, T. Heron, and W. Heward, *Applied Behavior Analysis*, 3rd ed. London: Pearson, 2019.

[31] W. Little, *Introduction to Sociology - 2nd Canadian Edition*, 2nd ed., 2016.

[32] D. D. J. C. D. Mitchell, *Communication Theory Today*. Stanford University Press, 2013.

[33] E. Goffman, *The Presentation of Self in Everyday Life*. New York: Anchor Books: A Division of Random House, 1959.

[34] J. Sternberg, *Misbehavior in Cyber Places: The Regulation of Online Conduct in Virtual Communities on the Internet*. Rowman & Littlefield, 1994.

[35] www.psychologytoday.com, *Face-to-Face Social Contact Reduces Risk of Depression*. Psychology Today Canada, 2022.

[36] E. T. Hall, *The Hidden Dimension*. Anchor Books, 1966.

[37] M. Newman, *Networks*. Oxford University Press, 2018.

[38] S. Brin and L. Page, "The anatomy of a large-scale hypertextual web search engine," *Computer networks and ISDN systems*, vol. 30, no. 1-7, pp. 107–117, 1998.

[39] R. I. Dunbar, "Neocortex size as a constraint on group size in primates," *Journal of human evolution*, vol. 22, no. 6, pp. 469–493, 1992.

[40] O. Perron, "Zur theorie der matrices," *Mathematische Annalen*, vol. 64, no. 2, pp. 248–263, 1907.

[41] G. Frobenius, "Über matrizen aus nicht negativen elementen," *Sitzungsberichte der Königlich Preussischen Akademie der Wissenschaften zu Berlin*, pp. 456–477, 1912.

[42] J. A. Ailshire and S. A. Burgard, "Family relationships and troubled sleep among us adults: examining the influences of contact frequency and relationship quality," *Journal of health and social behavior*, vol. 53, no. 2, pp. 248–262, 2012.

[43] M. Luo, B. Macdonald, and G. Hülür, "Not "the more the merrier"': Diminishing returns to daily face-to-face social interaction frequency for well-being in older age," *The Journals of Gerontology: Series B*, vol. 77, no. 8, pp. 1431–1441, 2022.

[44] D. W. Russell, "Ucla loneliness scale (version 3): Reliability, validity, and factor structure," *Journal of personality assessment*, vol. 66, no. 1, pp. 20–40, 1996.

[45] G. Hawthorne, "Measuring social isolation in older adults: development and initial validation of the friendship scale," *Social indicators research*, vol. 77, pp. 521–548, 2006.

[46] R. Wittek and R. Wielers, "Gossip in organizations," *Computational & Mathematical Organization Theory*, vol. 4, pp. 189–204, 1998.

[47] T. J. Grosser, V. Lopez-Kidwell, and G. Labianca, "A social network analysis of positive and negative gossip in organizational life," *Group & Organization Management*, vol. 35, no. 2, pp. 177–212, 2010.

[48] D. J. Watts and S. H. Strogatz, "Collective dynamics of 'small-world' networks," *Nature*, vol. 393, no. 6684, pp. 440–442, Jun 1998. [Online]. Available: https://doi.org/10.1038/30918

[49] S. Yin, Y. Hu, and Y. Ren, "The parallel computing of node centrality based on gpu," *Mathematical Biosciences and Engineering*, vol. 19, no. 3, pp. 2700–2719, 2022.

[50] M. E. J. Newman, "Assortative mixing in networks," *Physical Review Letters*, vol. 89, no. 20, oct 2002. [Online]. Available: https://doi.org/10.1103%2Fphysrevlett.89.208701

[51] J. T. Cacioppo, J. H. Fowler, and N. A. Christakis, "Alone in the crowd: the structure and spread of loneliness in a large social network." *Journal of personality and social psychology*, vol. 97, no. 6, p. 977, 2009.

[52] B. Lovell, "'we are a tight community': social groups and social identity in medical undergraduates," *Medical education*, vol. 49, no. 10, pp. 1016–1027, 2015.

[53] E. P. K. Tsang, *Foundations of constraint satisfaction*, ser. Computation in cognitive science. Academic Press, 1993.

[54] U. Brandes, "A faster algorithm for betweenness centrality," *Journal of mathematical sociology*, vol. 25, no. 2, pp. 163–177, 2001.

[55] M. Gast, *Building Applications with IBeacon: Proximity and Location Services with Bluetooth Low Energy*. O'Reilly Media, 2014.

[56] D.-Y. Kim, S.-H. Kim, D. Choi, and S.-H. Jin, "Accurate indoor proximity zone detection based on time window and frequency with bluetooth low energy," *Procedia Computer Science*, vol. 56, pp. 88–95, 2015.

[57] A. Montanari, S. Nawaz, C. Mascolo, and K. Sailer, "A study of bluetooth low energy performance for human proximity detection in the workplace," in *2017 IEEE International Conference on Pervasive Computing and Communications (PerCom)*. IEEE, 2017, pp. 90–99.

[58] A. Barrat, C. Cattuto, V. Colizza, J. Pinton, W. V. den Broeck, and A. Vespignani, "High resolution dynamical mapping of social interactions with active RFID," *CoRR*, vol. abs/0811.4170, 2008. [Online]. Available: http://arxiv.org/abs/0811.4170

[59] Țepordei AM, Z. AS, D.-G. LR, C.-T. I, M. C, S. D, and L. AV, "Children's peer relationships, well-being, and academic achievement: the mediating role of academic competence," *Frontiers in Psychology*, vol. 14, p. 1174127, 2003.

[60] 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.

[61] L. Isella, J. Stehlé, A. Barrat, C. Cattuto, J.-F. Pinton, and W. Van den Broeck, "What's in a crowd? analysis of face-to-face behavioral networks," *Journal of theoretical biology*, vol. 271, no. 1, pp. 166–180, 2011.

[62] 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.

[63] E. Colosi, G. Bassignana, D. A. Contreras, C. Poirier, P.-Y. Boëlle, S. Cauchemez, Y. Yazdanpanah, B. Lina, A. Fontanet, A. Barrat *et al.*, "Screening and vaccination against covid-19 to minimise school closure: a modelling study," *The Lancet Infectious Diseases*, vol. 22, no. 7, pp. 977–989, 2022.

[64] R. Mastrandrea, J. Fournet, and A. Barrat, "Contact patterns in a high school: a comparison between data collected using wearable sensors, contact diaries and friendship surveys," *PloS one*, vol. 10, no. 9, p. e0136497, 2015.

[65] M. C. Kiti, M. Tizzoni, T. M. Kinyanjui, D. C. Koech, P. K. Munywoki, M. Meriac, L. Cappa, A. Panisson, A. Barrat, C. Cattuto *et al.*, "Quantifying social contacts in a household setting of rural kenya using wearable proximity sensors," *EPJ data science*, vol. 5, pp. 1–21, 2016.

[66] D. O. Hasan, "Iot-based smart education: A systematic review of the state of the art," *Journal of Intelligent Systems and Information Technology*, vol. 2, no. 1, 2025.

[67] C. Thompson, A. Gordon, K. Khaliq, A. Daffu-O'Reilly, T. Willis, C. Noakes, and K. Spilsbury, "Quality in care homes: How wearable devices and social network analysis might help," *Plos one*, vol. 19, no. 5, p. e0302478, 2024.

[68] G. Koulouras, S. Katsoulis, and F. Zantalis, "Evolution of bluetooth technology: Ble in the iot ecosystem," *Sensors (Basel, Switzerland)*, vol. 25, no. 4, p. 996, 2025.

[69] C. A. Thompson, A. Daffu-O'Reilly, T. Willis, A. Gordon, C. Noakes, K. Khaliq, A. Farrin, A. Kemp, T. Hall, C. Bojke *et al.*, "'smart'ble wearables for digital contact tracing in care homes during the covid-19 pandemic—a process evaluation of the contact feasibility study," *Implementation Science Communications*, vol. 4, no. 1, p. 155, 2023.

[70] L. F. Berkman and T. Glass, "Social integration, social networks, social support, and health," in *Social Epidemiology*. Oxford University Press, 03 2000.

[71] K. M. Abbott, J. P. Bettger, K. Hampton, and H.-P. Kohler, "Exploring the use of social network analysis to measure social integration among older adults in assisted living," *Family & Community Health*, vol. 35, no. 4, pp. 322–333, 2012.

[72] M. Teston, E. Sturaro, E. Muñoz-Ulecia, A. Tenza-Peral, S. Raniolo, E. Pisani, C. Pachoud, M. Ramanzin, and D. Martín-Collado, "Participatory approaches and social network analysis to analyse the emergence of collective action for rural development: a case study in the spanish pyrenees," *Italian Journal of Animal Science*, vol. 23, no. 1, pp. 504–522, 2024.

[73] A. Garrote, C. L. Zurbriggen, and S. Schwab, "Friendship networks in inclusive elementary classrooms: Changes and stability related to students' gender and self-perceived social inclusion," *Social Psychology of Education*, vol. 26, no. 6, pp. 1479–1497, 2023.

[74] F. Karimi and P. Matous, "Mapping diversity and inclusion in student societies: A social network perspective," *Computers in Human Behavior*, vol. 88, pp. 184–194, 2018.

[75] N. S. Pearson, J. C. Major, A. Godwin, and A. Kirn, "Using social network analysis to study the social structures of inclusion," in *2018 ASEE Annual Conference & Exposition*, no. 10.18260/1-2–31211. Salt Lake City, Utah: ASEE Conferences, June 2018. [Online]. Available: https://peer.asee.org/31211

[76] S. S. Singh, S. Muhuri, S. Kumar, and J. Barua, "From nodes to knowledge: Exploring social network analysis in education," *ACM Transactions on the Web*, vol. 19, no. 1, pp. 1–36, 2025.

[77] V. Martínez, F. Berzal, and J. C. C. Talavera, "A survey of link prediction in complex networks," *ACM Comput. Surv.*, vol. 49, no. 4, pp. 69:1–69:33, 2017.

[78] Z. L. Li, X. Fang, and O. R. L. Sheng, "A survey of link recommendation for social networks: Methods, theoretical foundations, and future research directions," *ACM Trans. Manage. Inf. Syst.*, vol. 9, no. 1, Oct. 2017.

[79] M. C. Pham, Y. Cao, R. Klamma, and M. Jarke, "A clustering approach for collaborative filtering recommendation using social network analysis," *J. Univers. Comput. Sci.*, vol. 17, no. 4, pp. 583–604, 2011.

[80] G. Carullo, A. Castiglione, A. D. Santis, and F. Palmieri, "A triadic closure and homophily-based recommendation system for online social networks," *World Wide Web*, vol. 18, no. 6, pp. 1579–1601, 2015. [Online]. Available: https://doi.org/10.1007/s11280-015-0333-5

[81] A. Bashardoust, S. A. Friedler, C. Scheidegger, B. D. Sullivan, and S. Venkatasubramanian, "Reducing access disparities in networks using edge augmentation," in *Proceedings of the 2023 ACM Conference on Fairness, Accountability, and Transparency, FAccT 2023, Chicago, IL, USA, June 12-15, 2023*. ACM, 2023, pp. 1635–1651. [Online]. Available: https://doi.org/10.1145/3593013.3594105

[82] J. A. Fuhse and N. Gondal, "Networks from culture: Mechanisms of tie-formation follow institutionalized rules in social fields," *Social Networks*, vol. 77, pp. 43–54, 2022.

[83] R. Spears, "Social influence and group identity," *Annual review of psychology*, vol. 72, no. 1, pp. 367–390, 2021.

[84] M. Asif, M. Abrar, A. Salam, F. Amin, F. Ullah, S. Shah, and H. Al-Salman, "Intelligent two-phase dual authentication framework for internet of medical things," *Scientific Reports*, vol. 15, no. 1, p. 1760, 2025.

[85] C. J. Roseth, D. W. Johnson, and R. T. Johnson, "Promoting early adolescents' achievement and peer relationships: the effects of cooperative, competitive, and individualistic goal structures," *Psychological bulletin*, vol. 134, no. 2, pp. 223–246, 2008.

[86] H. Goswami, "Evaluation of a multi-component and multi-agent intervention to improve classroom social relationships among early elementary school-age children," *Social Indicators Research*, vol. 107, pp. 575–588, 2012.

[87] F. J. G. Bacete, G. Marande, and A. Y. Mikami, "Evaluation of a multi-component and multi-agent intervention to improve classroom social relationships among early elementary school-age children," *Journal of school psychology*, vol. 77, pp. 124–138, 2019.

[88] F. Valbusa, , and M. Ghio, "Potenziare le relazioni cooperative tra pari: una ricerca educativa in una scuola primaria," *RicercAzione*, vol. 15, no. 2, pp. 113–133, 2033.

[89] N. Klang, I. Olsson, J. Wilder, G. Lindqvist, N. Fohlin, and C. Nilholm, "A cooperative learning intervention to promote social inclusion in heterogeneous classrooms," *Frontiers in psychology*, vol. 11, p. 586489, 2020.

[90] D. W. Johnson and R. T. Johnson, "An educational psychology success story: Social interdependence theory and cooperative learning," *Educational researcher*, vol. 38, no. 5, pp. 365–379, 2009.

[91] K.-A. Allen, M. L. Kern, C. S. Rozec, D. M. McInerney, and G. M. Slavich, "Belonging: A review of conceptual issues, an integrative framework, and directions for future research," *Australian journal of psychology*, vol. 73, no. 1, pp. 87–102, 2021.

[92] J. A. Fredricks and J. S. Eccles, "Developmental benefits of extracurricular involvement: Do peer characteristics mediate the link between activities and youth outcomes?" *Journal of Youth and Adolescence*, vol. 34, no. 6, pp. 507–520, 2005.

[93] A. C. Fletcher and R. A. Shaw, "Sex differences in associations between parental behaviors and characteristics and adolescent social integration," *Social Development*, vol. 9, no. 2, pp. 133–148, 2000.