Exploring Space Syntax on Entrepreneurial Opportunities with Wi-Fi Analytics

Afra Mashhadi[†], Utku Günay Acer[†], Aidan Boran[†], Philipp M. Scholl[‡], Claudio Forlivesi[†], Geert Vanderhulst[†], Fahim Kawsar[†]
[†]Nokia Bell Labs, [‡]Albert-Ludwigs-Universität Freiburg

ABSTRACT
Industrial events and exhibitions play a powerful role in creating social relations amongst individuals and firms, enabling them to expand their social network so to acquire resources. However, often these events impose a spatial structure which impacts encounter opportunities. In this paper, we study the impact that the spatial configuration has on the formation of network relations. We designed, developed and deployed a Wi-Fi analytics solution comprising of wearable Wi-Fi badges and gateways in a large scale industrial exhibition event to study the spatio-temporal trajectories of the 2.5K+ attendees including two special groups: 34 investors and 27 entrepreneurs. Our results suggest that certain zones with designated functionalities play a key role in forming social ties across attendees and the different behavioural properties of investors and entrepreneurs can be explained through a spatial lens. Based on our findings we offer three concrete recommendations for future organisers of networking events.

ACM Classification Keywords
C.3. Computer Systems Organization: Special-Purpose and Application-Based Systems

Author Keywords
People Analytics; Space Syntax; Network Sensing

INTRODUCTION
With over 122 billion dollars of market value, the event industry such as conferences and summits provide a *setting* in which people from diverse organisations and with diverse purposes assemble to announce new products, transact business and most importantly establish new social relations [26]. It is through these network connections the entrepreneurs could identify opportunities and resources [1, 5] and establish trust on their firms [14].

The impact of these established connections, i.e. *social capital*, has been extensively studied in the field of organisational science. This body of literature has reported a direct relation between the social capital of entrepreneurs and their probability to succeed in cultivating resources. More recently, organisational theorists have examined the social structure and the formation of social ties during *events* and have shown that entrepreneurs who attended heterogeneous events were more likely to extend their networks through weak ties bridging structural gaps in their network [39]. Literature has also studied the impact of homophily [31] on cultivating social ties. Stam et al. have shown that homogenous events that induced homophily by inviting a specific set of participants led to reinforcement of already established ties by entrepreneurs [39].

Despite the recent focus on understanding the formation of social ties during the events, the impact of spatial characteristics in which these social activities are embedded has received little attention. Although network theories have mostly explored the relationship formation as a function of individuals' attributes, they have not considered the event setting that imposes a variety of structured components and constraints both in terms of program (e.g., structured presentations) and spatial configuration on the participants.

In this work, we thus examine *"what impact spatial layout has on the opportunities available to the entrepreneurs to strengthen their social ties?"* More precisely, we study the spatio-temporal behavioural properties of entrepreneurs and the source of their social capital (e.g., investors and other entrepreneurs) with an aim to uncover the extent to which their network formations are constrained by structural opportunities that are imposed by space syntax.

To this end, we designed and developed a Wi-Fi analytics solution comprising of wearable Wi-Fi badges and gateways. These gateways capture Wi-Fi signals radiating from wearable badges and anonymous Wi-Fi enabled devices to create spatio-temporal trajectories that can be studied to understand the relationships between spatial layout and behavioural properties of people. We had the unique opportunity to deploy this solution in a large scale industrial event and capture the spatio-temporal trajectories of 2.5K+ attendees including two special groups: 34 investors and 27 entrepreneurs[1]. The event attracted 40K+ people from 134 different countries and took place over three days in early November 2015 in a European city. The ±6000 sq. m. venue of the event was organised to create structural opportunities for the participants to meet and

[1]The consent to record the trajectory of those wearing the badges were given by the individuals themselves, and the consent to gather people analytics data from all the attendees were given by the event organisers.

UbiComp '16, September 12 - 16, 2016, Heidelberg, Germany

ACM 978-1-4503-4461-6/16/09…$15.00
DOI: http://dx.doi.org/10.1145/2971648.2971745

interact through technical presentation sessions, demo booths, marketing booths, startup pitch sessions as well as informal spaces. Collectively, these spaces afforded us to uncover the dynamics of the space syntax on entrepreneurial opportunities.

Our results show that spatio-temporal behavioural properties of different groups can be explained through a spatial lens. For instance, we observed that the investors followed a caveman mobility behaviour, i.e. they spent most of their time in a restricted access area designated to them, with sporadic short visits to other areas of the venue. In contrary, the entrepreneurs spent their time opportunistically visiting many different areas with occasional visits to their designated zone. We found that certain spaces with designated functionalities play a key role in forming social ties across attendees. Taken together these and the rest of our findings demonstrate the many ways in which space layout influences entrepreneurial opportunities. Based on these observations we recommend a set of guidelines for space design that induces informal social ties, such as to create multiple instances of the private areas to compel move between the spaces and to ensure that all the paths between these private areas pass tangent to the public areas, providing more exposure opportunities.

RELATED WORK

A rich body of literature in the past has studied crowd behaviour at large events. Location has always been the key signal in all these works, and typically the locations of the people are acquired through dedicated mobile applications [2, 33] or sensor tags [22, 24]. The former approach often uses either GPS or Wi-Fi or Bluetooth or Cellular signals or a combination of them to gather location information, whereas the latter approach in the literature exploited Bluetooth signals primarily.

For instance, in [22], Bluetooth tags have been distributed to participants of a scientific conference to study the mobility of the participants [22]. Similar tags are used by Jamil et al. to study the movement and community structure of pilgrimage during the religious festival of Hajj [24]. However, the ubiquity of mobile phones made it possible to use network signals to trace and track people to study mobility and behaviour without the need for specialised tags. Larsen et al. use Bluetooth signals emitted from smart phones to analyse the behaviour of participants in a large music festival [28]. Other works investigate the mobility and interactions of people by capturing Bluetooth signals from their devices with scanners [25, 36]. Similarly, in [41, 43], Bluetooth traces from smart phones are used to extract the trajectory of people during a large event. Bluetooth signals such as mean signal strength and the number of devices were also used to estimate the crowd density in public events [45]. To summarise, prior work on crowd analysis widely used short ranged radio technology such as Bluetooth and RFID [10] to collect information about participants. Though Wi-Fi signals have been widely leveraged for indoor localisation [21, 29, 46], Wi-Fi based systems have not been deployed to understand the behaviour of the crowd or the individuals.

Another body of work has focused on tracking face-to-face interactions and the impact of the space on enabling them [7, 42]. These works are often applied to controlled environments and settings such as offices. In this paper we aim to understand the role that space plays in large uncontrolled settings (such as events) instead of discovering interactions.

In the field of organisational sciences, researchers have paid close attention to entrepreneurial activities and their formed networks. These works commonly acknowledged that the opportunities entrepreneurss are exposed to leads to the establishment of ties which in turn results in resource acquisition by those entrepreneurs who are somehow better connected [47]. Studying entrepreneurs activities in the social settings such as parties and mixers are a more recent focus of this field. In this vein, researchers have examined entrepreneurs' interactions during the events through monitoring the content of the conversation or face-to-face interactions [23, 39, 40]. Although researchers have begun to explore the generalizability of cultivating network connections in different settings (e.g., heterogeneous vs homogenous events), it is still unclear what is the impact of the spatial configuration of the event venue on the expansion of social capital.

SYSTEM FOR LARGE SCALE PEOPLE ANALYTICS

Large scale people analytics in a semi-public place demands careful system design due to a number of challenges imposed by the nature of the place, dynamics of the crowd behaviour, and congested network conditions. In this section, we first discuss some of these design challenges that we have considered while building our system and then we present different components of the system and their operational principles.

Design Challenges and Decisions

The primary signals required for modelling crowd behaviour in a semi-public place are the spatio-temporal trajectories of people i.e., the main objective of the system is to estimate the positions of different individuals at different points in time to construct the trajectories. There are multiple alternatives to acquire these trajectories with varying characteristics in terms of system requirements, measurement complexity, energy constrains, participation scale and privacy concerns. For instance, past research on large scale people analytics used smartphone or wearable beacons with GPS [37] or Bluetooth [24, 41, 45] as key sensing modalities to extract spatio-temporal trajectories. Given our study was confined in an indoor space, GPS was excluded from our design alternatives. Though a Bluetooth based approach either with dedicated badges or smartphones is accurate in location estimates, it comes at the expense of (i) high deployment costs as many receivers are required due to the shorter communication range, (ii) limited participation especially with smart phones, as Bluetooth is turned off most of the times [9], and (iii) high energy cost of Bluetooth [19] and the fact that preserving smartphone battery is critical considering the scarcity of the electricity points in semi-public places. An alternative method is to use the cellular information of smartphones such as data packets or call detail records. Although this approach might yield in a larger participation, it requires specialised basestations and faces challenging network conditions in crowded public events. Besides, the range of cellular information is too wide (a few kilometres) to estimate the place-level granularity.

In this work, we diverted from these approaches and relied on Wi-Fi signals coupled with existing Wi-Fi infrastructure to estimate spatio-temporal trajectories of people. Basically, we captured Wi-Fi signals radiating from Wi-Fi enabled smart devices such as smart phones, tablets, wearables, etc., and badges by a set of pre-deployed Wi-Fi gateways to estimate the position of people and then to build their spatio-temporal trajectories. Our rationale behind using Wi-Fi signals is grounded upon the findings from past research suggesting that people carry their phones with them most of the time [27] having Wi-Fi switched on especially in places with free availability of the Wi-Fi network [9]. We argue that this is a reasonable assumption for our target place, e.g., an industrial exhibition event. Furthermore, Wi-Fi has a wider coverage than Bluetooth thus demands fewer receivers for capturing Wi-Fi signals. However, to achieve the goal of this work with our proposed approach, we needed to overcome two major challenges as discussed below:

Detection Reliability - We relied on Wi-Fi gateways to capture Wi-Fi signals originating from the users' devices. However, as pointed out in [34], no passive Wi-Fi tracking system can guarantee accurate tracking performance, due to a number of issues: (i) temporal and spatial sparsity of detections across gateways, (ii) unpredictable path loss, e.g., caused by walls, furniture, people, etc. and (iii) heavily congested communication channels. To address these issues we have taken two important design decisions: firstly, we leveraged Wi-Fi *probe requests* instead of data packets to estimate positions of the people. A probe request - further referred to as *probe* - is a Wi-Fi management frame that is either directed to a specific network or broadcasted to any network within range of a mobile device while it searches for a known Wi-Fi network to join. This probe based approach allowed us to cope with congested communication channels. Secondly, we deployed the gateways as isolated network nodes with no external connectivity to further minimise packet loss, and reduce congestion. Essentially these gateways captured probes that were originated from mobile devices and stored them locally.

Guaranteed Participation - In this work, we intended to study the behavioural traits of specific groups of people, e.g., entrepreneurs and investors. As such, we decided to use dedicated Wi-Fi badges that were distributed to a subset of these groups and designed with specific operational principles. These badges emit probes systematically so that nearby Wi-Fi gateways can capture them, which are then used to estimate the positions of the wearers in the venue. Since our gateways could also capture probes originating from anonymous smart devices, we also had the opportunity to study the overall crowd dynamics in the event. This two-tier approach guaranteed us to acquire the spatio-temporal trajectories of the entrepreneurs and investors at a fine granularity while still capturing the anonymous trajectories of the crowd with a reasonable coverage.

Based on the design decisions described above, we developed a system comprising of two components, a wearable Wi-Fi badge and a set of Wi-Fi gateways.

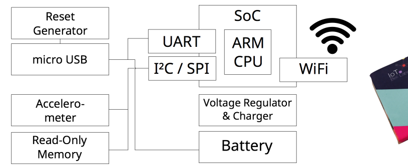


Figure 1: (a) The schematic design and (b) the final badge.

Wearable Wi-Fi Badge

We designed a credit card sized wearable badge that emits probes systematically. In the following, we describe its different components and working principles.

Hardware Component

Two concerns mainly drove the hardware design of the badge: physical size and functional requirements. Physically the badge should equal the size of a standard credit card (including battery, electronics and all outside connections) to ensure that it could be attached to a standard conference lanyard. Functionally, reprogramming and re-charging of the battery should be achieved by the user. Taking both these concerns into account, we designed a badge that was implemented using a custom designed PCB with a dimension of 85mm x 55mm x 2mm. The badge was composed of an ESP8266[2], a SoC consisting of an ARM-based CPU and a 2.4GHz Wi-Fi controller, a Freescale MMA8452Q accelerometer and associated power regulation and battery charging circuitry. We opted for a ultra thin 180mAH LiPo battery which was regulated to provide the system with power. The accelerometer was used to wake up the system from sleep on motion. The badge can be charged and programmed using a micro USB connector. Figure 1 shows the schematic of the badge.

Software Component

We developed an energy aware firmware to broadcast the 802.11 management probes systematically. The firmware implemented a simple algorithm that could detect when the badge was in motion for multiple seconds using the accelerometer and used this as a trigger to send probes. If no motion was detected for a specific period of time, then the badge followed a pre-defined schedule for sending probes (one per minute). The probe functionalities, i.e., custom header construction, emit rate, etc. were implemented in C on top of a modified Wi-Fi stack of ESP firmware. Specifically, we control the `Sequence Control` field of the probe frame to uniquely detect a probe emitted from a specific badge (identified by its MAC address). The motion detector was implemented in C using the standard I^2C library.

Energy Management

The ultra thin 180mAH battery had a total capacity to send 1840 probes. Past research has shown that Wi-Fi packet transmission is energy expensive [8]. To ensure that a fully charged badge could operate the whole duration of the events, we used a timer that put the badge into deep sleep after 10 operational hours (08:00 - 18:00), and woke up the badge after 14 hours. During the operational cycle the badge followed the algorithm mentioned earlier for sending probes. This approach ensured

[2]http://www.esp8266.com

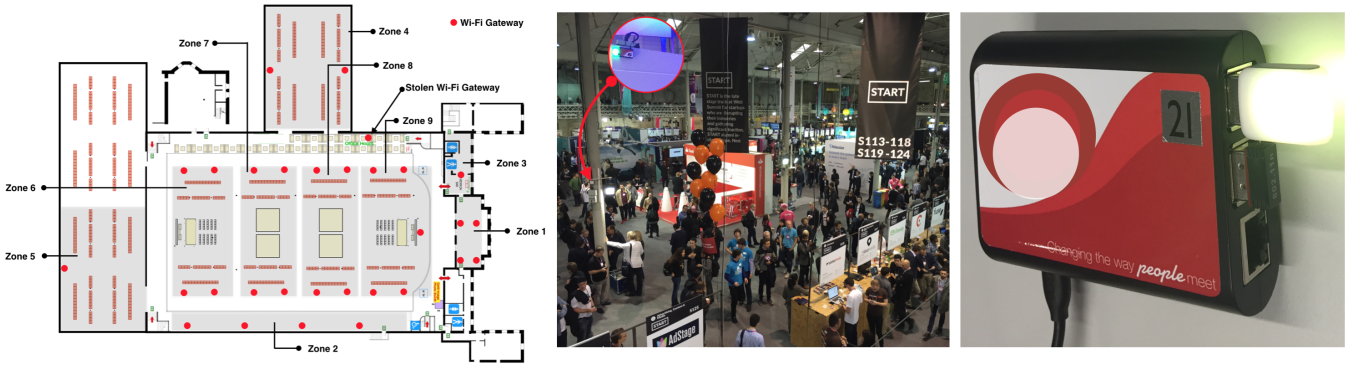


Figure 2: (a) Venue floor plan and deployment map, (b) snapshot of the event and (c) the Wi-Fi gateway.

that we capture probes for critical times compared to the simpler approach of sending probes at regular intervals until the battery is drained. This design also enabled the badges to operate effectively over the 3 days of the event without the need for recharging.

Wi-Fi Gateway

In our system, Wi-Fi gateways captured the probes emitted from wearable badges and other smart devices . These gateways were implemented using Raspberry PI 2 Model B one board computers equipped with a Wi-Fi dongle (Ralink 5370) and a programmable LED lamp. The probe capturing functionality was implemented in C++ using the *libtins* library[3]. The LED lamp acted as an indicator that a gateway was operating normally. Figure 2(c) illustrates a gateway and its constituent components.

Each probe is assembled following the frame format of the IEEE 802.11 standard. Once a probe is captured, the gateway analyses its IEEE MAC header and Radiotap header to extract {`mac`, `sequence`, `RSSI`, `timestamp`} and stores these values for the probe in its storage. The `RSSI` field indicates the signal strength of the received probe at the gateway, whereas the rest of the fields uniquely identify the probe.

When Wi-Fi devices scan for networks, they send probes on every Wi-Fi channel (2.4GHz or 5GHz). Given the Wi-Fi infrastructure of the event was on the 5GHz band with each access point working on a different channel, we set the Wi-Fi dongle in monitor mode on 2.4GHz band to capture the probes emitted from the badges and smart devices. Although probes are emitted less frequently than data frames by a smartphone, given the crowded nature of the event and congested open Wi-Fi network, we expected that devices would often lose the connection, or look for a better connection resulting in the transmission of probes. Even if the device keeps the connection to a particular access point, it still scans Wi-Fi networks with scanning interval defined by the vendor [17] that allows device to be detected by the deployed gateways.

DEPLOYMENT

Our study was conducted in early November, 2015 in a three-day industrial exhibition in western Europe. The event attracted over 40K attendees including representatives from major IT companies, small to medium sized startups, media companies, venture capital firms, etc. The event had many technical presentation sessions, demo booths, marketing booths, and startup pitch sessions. Besides regular attendees, and presenters, there were two special attendee groups: entrepreneurs and investors. The $\pm$6000 sq. m. venue was organised strategically to create multiple opportunities for these two groups to meet, interact, and socialise. We deployed 30 gateways in 9 different zones. Zone 2 was restricted to investors only, however, they could bring entrepreneurs or other visitor along with them inside the lounge. Zone 1 and zone 3 were restricted to investors and entrepreneurs. The rest of the zones had no access control and were primarily organised to house the demo booths, and two startup pitch stands. Table 1 describes the functionality of each zone along with their access right. These zones and the location of the deployed gateways and a snapshot from the exhibition are shown in Figure 2(a,b). Out of the 30 gateways, we were able to retrieve 29 gateways after the event closing. We distributed 85 badges on the first day of the event to a set of pre-designated entrepreneurs and investors who had given consent to take part in this study. However, 24 of them were observed only for a very short time and thereby excluded from our analysis. Out of the remaining 61 badges, 34 of them were worn by the investors, and the rest 27 were worn by the entrepreneurs during the event.

Zone	Functionality	Access
1	Starter's Lounge	Entrepreneurs & Investors
2	Investors' Lounge	Investors
3	Coffee Bar	Entrepreneurs & Investors
4,5,7,8	Exhibition Booth	Everyone
6,9	Stages	Everyone

Table 1: Different zones of the venue and their access policies.

DATA DESCRIPTION AND PRE-PROCESSING

We begin by describing the crowd data that was captured by our system during the event. As we discussed in the earlier section, all captured probes either from wearable badges or anonymous smart devices were locally stored in the Wi-Fi gateways. After the event, we collected the stored logs from all 29 gateways, merged them, and finally sorted them by timestamp to create a large trace dataset ordered by time containing all the probes.

[3] http://libtins.github.io

Data Description

All together, 29 gateways captured 7M+ probes from over 290K unique MAC addresses including 69K+ probes from 85 badges. However, a large portion of these records did not contain meaningful information as most of the MAC addresses were seen only a few times or over a very short period of time. 38 % of the MAC addresses seen in the traces had only a single probe and more than 85% of the MAC addresses had less than 10 probes in total.

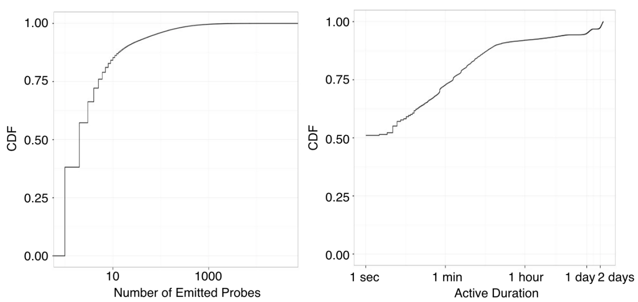


Figure 3: Distribution of the emitted probes and active duration for each device.

As illustrated in Figure 3 (a), the distribution of probes per device had a very long tail with a small number of non-mobile devices were seen sending over 40K probes. We confirm that these were the networked gateways that offered Wi-Fi connectivity, or P2P devices that searched constantly for peers. As depicted in Figure 3 (b) about 88% of the MAC addresses were seen only for a short duration (less than 10 minutes). In order to gain better insight from the traces, we decided to include only those MAC addresses that sent at least 300 and at most 1000 probes and were active for at least two different days in the event.

After this filtering, we ended up with 2526 MAC addresses (1.3M probes), 61 badges (56K probes) yielding two datasets: a crowd dataset that we have used subsequently for crowd analytics, and a badge dataset that we have used for analysing the behaviour of investors and entrepreneurs. Please note that in the ideal case we should have received 112240 probes from these 61 badges[4] if their wearers were in the venue all the time (08:00 - 18:00, all 3 days). Although possible, given the profile of these investors and entrepreneurs it is reasonable to assume that they did not spend the whole duration in the venue. Hence, we argue that with our coverage of 50.4% probes we would be able to uncover an illuminating picture of their behavioural dynamics during this event.

The summary of the entire and the filtered traces are given in Table 2.

	# MAC	# Probes	# Badges	# Badge Probes
Total	290740	7054970	85	69191
Filtered	2526	1325364	61	56615

Table 2: Description of the orginal and filtered datasets.

[4]Recall that each badge could send atmost 1840 probes with a fully charged battery until the battery is drained.

From Probe to Spatio-Temporal Trajectory

RSSI acquired from a probe indicates a symbolic position of a mobile device or a badge relative to a reference Wi-Fi gateway that captured the probe. This coarse grained localisation offering relative location instead of absolute location is known as proximity ranging [21, 29, 46] and is used as the primary technique in this work to construct spatio-temporal trajectories from probes. In proximity ranging, when more than one Wi-Fi gateway detects a device (which is the case in this work), it is considered to be co-located with the one that receives the strongest signal. However, simply selecting the strongest signal does not necessarily yield a reliable ranging performance due to the non-deterministic signal propagation characteristics of Wi-Fi signals. Besides, in our deployment we had multiple Wi-Fi gateways representing a zone.

We formulated this proximity ranging task as a multi-class classification problem by dividing our space in $K = 9$ non-overlapping spatial zones (see Table 1 and Figure 2(a)). For each device we used a one minute window ($w_d = 1$) to combine all the probes that were seen by different gateways (including the same probe captured by multiple gateways). For those gateways that did not capture the probe we used -100 db as RSSI. We constructed a feature vector containing 87 elements representing a device's position data within w_d. The first 58 elements were the mean and standard deviation of the RSSI received by 29 gateways, and the last 29 elements represented the frequency of each gateway within w_d. For training our classifier we collected ground truth data on the eve of the first day of the event. 5 persons wearing the badges that sent probes every 3 seconds spent 10 minutes at different positions in each zone. We recorded the closest zone manually as ground truth and annotated the records with start and end time of the dwell period in each zone. This phase yielded 50 minutes of data containing 1000 probes for each zone coming from 5 badges. We tested with three classifiers: Random Forest [6], Boosting [18] and Support Vector Machine (SVM) and their One-Vs-The-Rest and One-Vs-One variants. We found that SVM leveraging One-Vs-One classification scheme [38] performed the best with our training dataset. One-Vs-One classification uses an underlying binary classifier to compare every pair of classes and followed by a majority voting scheme to decide the winning class. For the SVM we have used a universal Gaussian kernel.

$$k(\mathbf{x}_i, \mathbf{x}_j) = \exp(-||\mathbf{x}_i - \mathbf{x}_j||^2/2/\sigma^2) \qquad (1)$$

where $\mathbf{x}_i$ represents the feature vector of a device for w_d.

To fix the hyper parameters of the SVM, namely σ in equation 1, we have used a validation procedure, in which we used 80% of the data for training and 20% for testing. With this classification scheme we have achieved a F_1 score of 0.92.

We processed the entire probe trace through this classifier to construct a minute by minute (when data was available) spatio-temporal trajectory for each device of the crowd dataset and badge dataset respectively. These trajectories were then used for subsequent analysis that we report next.

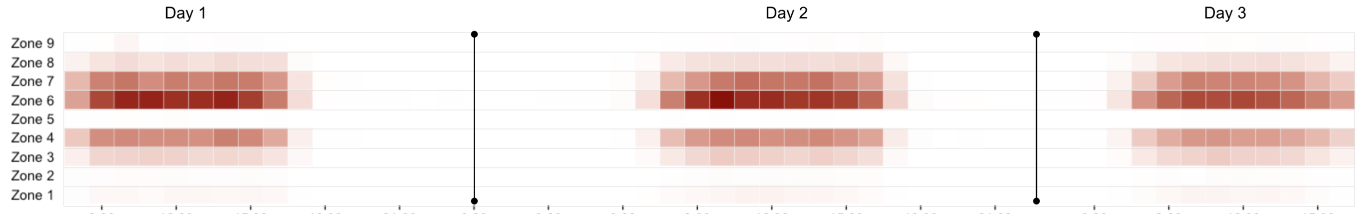


Figure 4: Spatio-temporal heat map of the crowd as captured by 29 gateways throughout the 3-day event, with colour intensity corresponding to the crowd density in different zones in different days.

RESULTS

The two datasets and their combination gave us a fascinating picture of the behavioural dynamics of the attendees, and the role that the spatial structure of the venue played in creating opportunities for social ties among the attendees. In this section, we discuss the study results from three perspectives. First, we discuss the crowd behaviour by analysing the crowd dataset. Then we discuss the space syntax by combining both datasets, and finally we discuss the behavioural dynamics of the entrepreneurs and investors by analysis of the badge dataset.

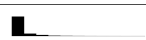
Property	μ	Distribution	σ
Duration in event (mins)	149.17		83.12
No. of visited zones	5.05		1.08
Duration per zone (mins)	29.52		47.46

Table 3: Statistics from the crowd dataset.

Understanding Crowd Dynamics

Table 3 describes the overall statistics of the behavioural properties of the crowd. We observe normal distributions for the number of visited zones, and the total duration that individuals spent in the event. Furthermore, we observe a long tail distribution on the duration spent in each zone. This suggests that a large proportion of the attendees were concentrated in only a few zones. Given the scale of the event that we studied, these results are not surprising as they reflect the natural behaviour that we expect from the attendees in an industrial event.

Figure 4 illustrates the heat map of the spatio-temporal density of the crowd throughout the event across 9 different zones.We notice that a set of zones, e.g., Zone 4, Zone 6 and Zone 7 were constantly crowded throughout the event. These three zones hosted the demo booths and one of the two pitch stages. This characteristic, however, is not present in other zones with similar functionalities, e.g., Zone 8 or Zone 9. Especially, Zone 9 had the second pitch stage. However to our surprise, we do not notice this zone as popular as its counterpart, i.e., Zone 6. We also observe here that Zone 5 had very little crowd footprint. We conclude that this low coverage was due to the fact that it was partially covered by only one gateway (see Figure 2(a)) and that the gateway was placed on one of the exit gates.

Next, we observe that Zone 1 and Zone 2 had relatively low density throughout the event. This is expected as these zones had access restriction and were only open to investors and entrepreneurs. Interestingly, Zone 3 which also had a restricted access policy did attract a relatively large number of attendees. Recall that this zone hosted a recreational Coffee Bar. These last observations have interesting implications that we elaborate later in the paper.

Understanding Space Dynamics

One of the key aspects that we want to explore in this work is the linkage that exists between spatial syntax and social ties observed in an industrial event. In particular, we want to answer whether spatial configuration plays a pivotal role in creating opportunities for different groups of attendees (visitors, investors, entrepreneurs, etc.) to bond socially that might lead to mutually beneficial working relationship in the future. In order to examine this aspect we employ the idea of *affiliation networks* [4, 16].

To construct the network, we combine both our datasets, and create a bipartite graph with two sets of nodes, S and F. S represents all the attendees, while F is the set of nodes that represent the zones (affiliations or foci) in which attendees spend their time during the event. An edge $\{(s,f) : s \in S \wedge f \in F\}$ exists iff s has spent at least 5 minutes in the focus zone f in at least one occasion, i.e., an attendee s is said to be affiliated with zone f if s/he has spent at least one contiguous 5 minute period throughout the 3-day event. In an affiliation network, when the members of S have social ties, the network transforms into a social-affiliation network, using which we can analyse the co-evolution of both the social and affiliation network. A new social tie might be formed due to a common attendee (triadic closure) or due to a common affiliation, i.e., dwell time in a zone (focal closure). Given we are interested in space dynamics, we primarily look at the focal closure which is an artefact of selection, i.e., forming links to others in zones with shared interest.

We define two metrics - *degree distribution* and *entropy* for the zones (i.e., nodes of the set F). The degree of an affiliation-node l (i.e., zone) is the number of distinct attendees that have visited and the affiliation entropy is essentially a measure of diversity of the zone [11]. If $P_l(u)$ is the fraction of visit in an affiliation l by attendee u, then the entropy of l is defined by:

$$Entropy(l) = -\sum_{u \in U_l} P_l(u) log P_l(u) \qquad (2)$$

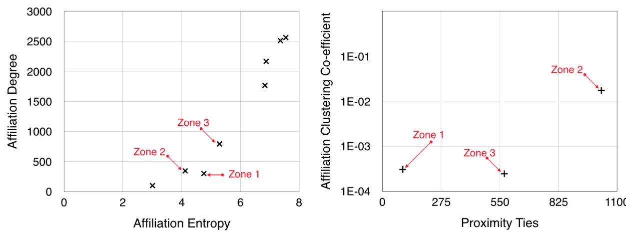


Figure 5: (a) Affiliation degree as a function of affiliation entropy (b) the relationship between clustering co-efficient and proximity ties of affiliation zones.

When the entropy of a zone is high, a large number of people visits the place in equal but small proportions. On the contrary when mass of $P_l(u)$ is concentrated around a small number of people, the diversity in the zone is low, so is the entropy. We first examine this aspect by correlating degree of an affiliation with its entropy. Not surprisingly we observed a strong correlation (Pearson's $r = 0.94$ and $p < 0.001$), i.e., as entropy increases so does the degree of a zone as shown in Figure 5(a). To further examine how affiliation entropy influences formation of new social bonds, we calculated pair-wise *proximity ties* across all the attendees in different zones. Here, we assume that a pair of attendees form a proximity tie if they spend at least 5 minutes in a shared zone simultaneously. We observe strong co-relation between the entropy of an affiliation zone and the volume of proximity ties observed in that zone (Pearson's $r = 0.71$ and $p < 0.001$). Note that we cannot claim that this result indicates that a social tie was formed between a pair due to their proximity tie and we acknowledge that more diversity implies random co-location. However, one should expect that the more common proximity ties two people have, the more probable it is for them to meet, and socially interact.

Further examining this, we made an interesting observation, for Zone 1, Zone 2 and Zone 3 we found a moderately negative correlation between their entropy and the volume of proximity ties observed in those zones (Pearson's $r = -0.67$ and $p = 0.091$). This has an interesting implication, as it suggests that as the degree and entropy of these affiliation nodes decrease they afford more opportunities for proximity ties, i.e., venture socialisation. Recall that all three zones had access restricted to only investors and entrepreneurs. To uncover how tightly these zones afford socialisation opportunities, we measure the *clustering co-efficient* of these zones. For a zone l with a degree $k > 1$, all possible social links between the attendees affiliated with l is $\frac{k(k-1)}{2}$. Given we have n proximity ties representing social links in that zone, we can compute the clustering co-efficient (CC) of l as:

$$CC(l) = \frac{n}{k(k-1)/2} \tag{3}$$

Essentially, $CC(l)$ expresses how closely connected are the people that visit the affiliation-node l. A higher $CC(l)$ indicates that people affiliated with l are tightly connected. As shown in Figure 5(b), while examining the clustering co-efficient of Zone 1, 2, and 3 as a function of their proximity ties, we clearly observe that as a zone exhibits higher clustering co-efficient it ventures more proximity ties. These results indicate that these three zones had interesting dynamics in affording social ties across the attendees. In the next section, we discuss these dynamics in more detail.

Understanding Group Dynamics

In this section we present the results and observations from the analysis of the data recorded through the badges. Table 4 presents the overall statistics of the behavioural properties observed from the 61 badges.

Property	μ	Distribution	σ
Duration in event (mins)	390.57		176.23
No. of visited zones	4.93		1.80
Duration per zone (mins)	7.53		12.08

Table 4: Statistics from the badges dataset.

Based on these statistics we could observe that there is a long tail distribution of the amount of time spent per area. In other words, there are many records with short stay in an area. We choose to filter these records. The analysis reported in the rest of this section is performed over the filtered data.

Understanding Dwelling Behaviour

First we look at how the investors and entrepreneurs spent their time in terms of spatial footprint. Figure 6 presents the total time spent in each zone by the investors and the frequency distribution of the number of distinct zones that were visited. Figure 7 depicts the same information for the entrepreneurs. As we can see the investors spent most of their time in zone 2, which corresponds to the Investors' Lounge and spent less dwelling time in other zones such as the stages and exhibition booths. Figure 8(a) shows that almost half of the investors spent more than 50% of their total time in the event in the Investor's Lounge (μ=0.63,σ=0.21). This figure indicates that on contrary to the investors, the entrepreneurs spent proportionally less time in their designated lounge.

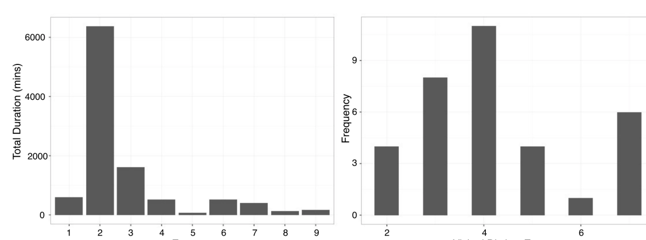


Figure 6: (a) Total duration (in min) investors spent in each zone, (b) Frequency distribution of the number of distinct zones they visited.

Moreover, from the frequency distribution of Figure 6(b) we can see that the majority of the investors visited less than 4 zones during the event. Comparing this behaviour with the entrepreneurs (Figure 7), we find that the entrepreneurs were significantly more likely to visit more number of zones in total (t=3.53, p-value=0.0008). Indeed they visited on average more than 5 zones (μ=5.81, σ=1.8).

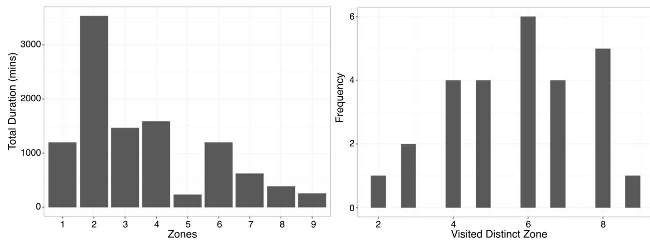


Figure 7: (a) Total duration (in min) entrepreneurs spent in each zone, (b) Frequency distribution of the number of distinct zones they visited.

To further investigate participants' dwelling behaviour in designated areas, we model participants' trajectories during the event as a directed graph $G<V,E>$ where the vertices are the zones and the edges correspond to the trajectories between the zones. We associate each edge with a weight w corresponding to the number of times that edges was traversed by the same participant. We define two metrics: *degree* and *average path length*.

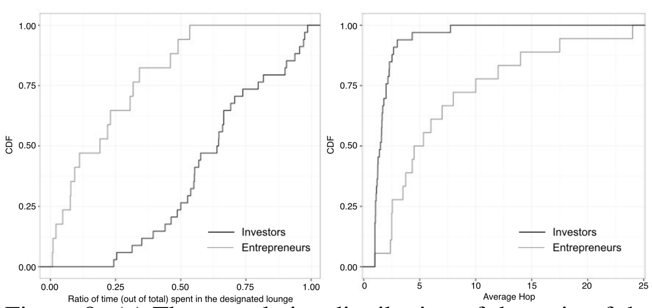


Figure 8: (a) The cumulative distribution of the ratio of the time spent in the investors lounge. (b) Average Path Length Trajectory between base visits.

Degree: The degree of a vertex (i.e., zone) is the number of edges to/from other zones. We calculate the *out-flow* and *in-flow* degree from the *Investors' Lounge* and *Starter's Lounge*. First reporting on the mean degree of the Investors' Lounge based on the trajectory graphs of the investors, we observe out-flow degree of $\mu = 2.30$ ($\sigma = 0.76$) and an in-flow degree of $\mu = 2.60$ ($\sigma = 0.78$). Comparing this trajectory behaviour with those of the entrepreneurs from the Starter's Lounge, we observe a significantly lower out-flow degree $\mu = 0.88$ ($\sigma = 1.25$) and in-flow degree $\mu = 0.92$ ($\sigma = 0.91$). These numbers indicate that the investors left the Investors' Lounge in order to visit other zones and return back to the Investors' Lounge. This behaviour was significantly different from those of the entrepreneurs who exhibited lower out-flow degree ($t = 5.13, p - value < 0.001$). We also observe a similar in-flow characteristic in investors behaviour (and lack thereof in entrepreneurs), that is the investors returned to the Investors' Lounge from the different areas ($t = 7.51, p - value < 0.001$).

Average Path Length: Although the former metric allows us to understand the distribution of the edges to/from each zone, it does not capture any information regarding the trajectory of the participants once they leave the lounges. As we are interested to discover properties of the participants trajectories outside the lounge area, we also calculate the average path length between the visits to the lounge (i.e., investors lounge for the investors and start-up lounge for entrepreneurs). That is for each participant we measure the number of zones they visited once left and up to the time of their return to the corresponding lounge. We then compute the average of these path lengths per participant. Figure 8(b) illustrates this metric for both the investors and entrepreneurs. As it can be observed, the investors exhibit shorter average path length, where 80% of them only visit 2 zones before returning to the investors lounge. This is in contrast with the entrepreneurs who take much longer journeys out of their base area to explore opportunities.

From these results we can infer that the investors followed a mobility behaviour similar to so called Caveman Model [44] where there is an observed high stay time in their primary social clusters (i.e., investors lounge) and short sporadic visits to other areas with a low average path length. In studying trajectory features, we also observed strong behavioural homophily amongst the investors. Indeed, if we build *eigenbehaviours* [13] by applying principle component analysis to a matrix of the aforementioned features of the investors, we observe that 95% variance of our subjects' behaviour (investors) could be explained using only two principle components. These two components are the ones that have the largest eigenvalues for the average time spent outside of the lounge and the out-degree from the Investors' Lounge. Finally, although we cannot reason on the activities and interactions that took place inside the investors lounge, we believe our findings resonate with those of organisational science literature regarding the induced homophily amongst those of the same social status and income level [32].

Understanding Exposure and Interaction

The second dimension we study is the behaviour of entrepreneurs in regards with the opportunities offered to them by investigating the composition and properties of different spatial areas they explored during the event. In order to quantify the opportunities available to entrepreneurs, we borrow tenets from the social science and measure group exposure. Introduced first by Blau [3], *index of isolation* is a commonly used measure that quantifies the extent to which minority groups are exposed only to one another, rather than to non-minority groups. When this index is high, the communities suffer from segregation, limiting the chances of exposure to the alternative population. This index is computed as the minority-weighted average of each section's minority population. In the context of this work, this index is formulated as:

$$Index\ of\ isolation = \sum_{n=1}^{K} \left(\frac{ent_{i,t}}{ent_{total}} \cdot \frac{ent_{i,t}}{ent_{i,t} + inv_{i,t}}\right) \quad (4)$$

where K is the number of spatial areas, $ent_{i,t}$ and $inv_{i,t}$ are thesnumber of entrepreneurs and investors respectively in area i at time t and ent_{total} is the total number of entrepreneurs in event.

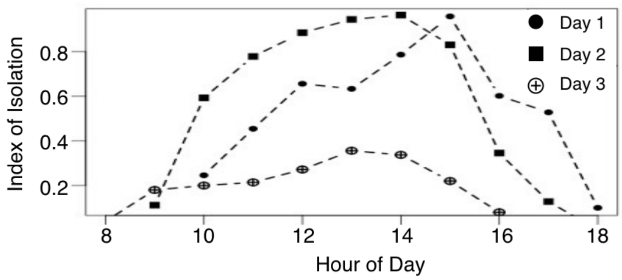


Figure 9: The index of isolation of the entrepreneurs.

Figure 9 illustrates the hourly snapshots of the *index of isolation* of the entrepreneurs over three days of the event. As it can be seen during the first and the second day the index of isolation is significantly high presenting a low exposure probability of entrepreneurs with the investors. We observe a much lower isolation index during the last day of the event which suggests that there were considerably higher exposures between the two groups on the final day.

To delve into this finding further, we study the spatial distribution of co-locations between entrepreneurs and investors. That is we measure the number of times a pair of an entrepreneur and an investor were co-present in any given location. Figure 10 presents the number of these co-located pairs for areas of the hall and during the three days lifespan of the event. As it can be seen during the first two days the two groups were co-located together in many different areas of the hall. On the third day however the co-locations mostly took place only in the Zone 3 and 2 corresponding to the Coffee Bar and the Investor's Lounge, thus providing a higher chances of exposure between the two groups. This figure also reports of a surprising insight, that is the high number of co-located pairs in zone 2[5]. Looking deeper at this insight, our further linear regression analysis indicates that the entrepreneurs who spent more time in the Coffee Bar area in proximity of the investors, were more likely to meet those investors in the Investor's Lounge ($r = 0.25, p - value < 0.01$).

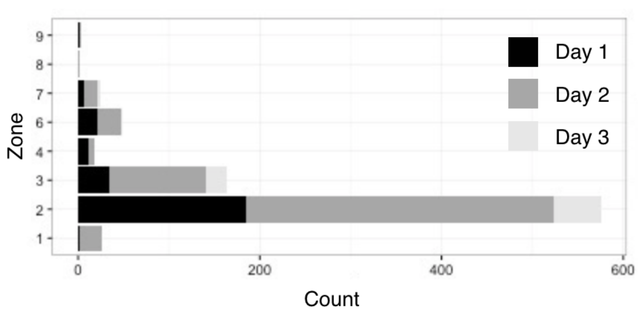


Figure 10: The total number of co-located pairs at different areas of the hall.

These results could be interpreted in two ways. Firstly, the concentration of the co-location pairs at the Investor's Lounge could suggest of pre-existing social ties between some entrepreneurs and investors. That is the individuals used the event to pre-scheduled meetings and mature their existing relationships. Secondly, it could be interpreted in the light of spatial characteristics and opportunities that different areas of the hall offer in enabling 'elemental encounters' [20]. In particular we observed significant correlation between the social proximity of pairs in the Bar area and the Investor's lounge. Although we cannot interpret this relation as a causation, we believe that the lower density and the informal setting of the Bar afforded both groups opportunities to encounter each other in a more meaningful way.

Research literature has shown that the accumulation of social capital [35] leads to an effect of "rich gets richer" where those with already established social status are offered opportunities that lead to cultivating more resources. Our results also resonated with those past findings, as we observed that the entrepreneurs who had already established themselves were invited to have access to restricted quieter areas such as Coffee Bar and Lounges. These areas afforded exposure opportunities for elemental encounter in an informal setting, enabling the entrepreneurs to further expand their social capital [35]. Based on these observations, we believe in such events it is possible for the social capital to substantially outperform the impact of human capital (e.g., education and personality traits) [12] in cultivating new relations. Our observation also suggests that despite the high financial cost that young entrepreneurial firms incur to participate in this type of events, they are perhaps the least beneficiary due to the spatial constraints that are imposed on them.

CONCLUDING REMARKS

In this paper we investigated the relationship between spatial layout of a large scale industrial event and the behavioural dynamics of its attendees. In particular we studied two special groups of attendees: 34 investors and 27 entrepreneurs to understand their spatio-temporal behaviour in the event. To this end, we designed, developed and deployed a Wi-Fi analytics solution comprising of wearable Wi-Fi badges and gateways in a large scale industrial exhibition event. The gateways captured Wi-Fi signals radiating from badges and anonymous smart devices to create a corpus of spatio-temporal trajectories of the 2.5K+ attendees. Borrowing tenets from affiliation network, organisation science, and social science literatures we analysed these trajectories to uncover a number of interesting insights. We observe that certain zones with designated functionalities play key role in forming social ties across attendees, and indeed spatio-temporal behavioural properties of different groups can be explained through a spatial lens. We found that two special groups (investors and entrepreneurs) behave very differently from each other. Taken together these and the rest of our findings can be translated into following implications and design recommendations for UbiComp research, and physical space management.

Decentralised System Design for People and Space Analytics: We have demonstrated the use of network sensing technology to model people's movement and provide insights on how socialisation unfolds in a physical space. We consider our systematic exploration of the design space, and the deep description of some of the core system components will enable future system designers to easily replicate our design

[5] Recall that the Investors Lounge (zone 2) had restricted access and was accessible to the entrepreneurs only if they were accompanied by an investor.

principles to build systems for large scale people and space analytics, especially in indoor settings, e.g., shopping malls, railway stations, enterprises etc. In addition, the Wi-Fi sensing system presented in this work can be applied to a diverse set of analytics applications beyond the ones discussed in this paper, e.g., understanding community structure, formation, and evolution, mobility modelling, etc. Furthermore, a Wi-Fi sensing system could be exploited to implement emerging RF-based social sensing such as *interactions* detection and *activity* detection [15] at a large scale.

Metrics for People Analytics: Many past research has explored spatio-temporal trajectories of people's movement primarily through a mobility lens. While we also investigated people's mobility patterns, one of the key aspects of our research was to understand the impact of space syntax on social behaviour. In so doing, we delved into the literature of related fields including organisation and social science to carefully define a set of metrics to analyse the data collected. Metrics such as *index of isolation*, *clustering co-efficient*, etc. are well understood and applied in their respective fields. In this work, we have shown that how these concepts could be applied to analyse people's movement data to extract spatio-social dynamics. We expect, these metrics and their applications will serve as a useful reference for other UbiComp researchers who may want to explore large scale people analytics.

Novel People Analytics Applications: We have shown in this work how to systematically transform ubiquitous Wi-Fi signals into social signals. These signals can easily be processed to translate into design strategies for new applications. Revisiting our industrial event scenario, event attendees today often rely on social networks (e.g., Twitter) to monitor the status, interactivity and atmosphere of different sessions during an event. A system such as ours could be used to visualise and provide feedback to the attendees on these social aspects. Certainly, the applications of our system are not limited to those of event industry and could be applied in various settings. For example, in a workplace setting our system could be deployed to understand the relationship between space syntax and interactions amongst the employees. It could be also used to promote collaborations and interactions amongst team members by capturing and visualising their interaction patterns [30]. We call attention to future application developers to carefully consider service design opportunities driven by people analytics. We hope that the system and insights offered in this work would assist their design process substantially.

People Analytics Driven Space Management: In this work, we primarily studied the impact of space syntax on the social behaviour of people in a public setting, and in particular for two specific groups, e.g., investors and entrepreneurs. Our results clearly indicate that we can infer the characteristics of spaces and their affordance towards social interaction. Based on these observations, we recommend space (be it an informal social gathering, or an industrial event, or a conference, or an enterprise, etc.) managers to cater for *intimacy gradient*. For instance, one way of incorporating these patterns into the design of the space is to ensure that all the paths between the private areas (e.g., lounges) pass tangent to the public areas which are the heart of the activities. We observed that how people with different profiles (e.g., investors) follow the concept of induced homophily where they stay at their designated zone (Investors' Lounge) with short sporadic visits to other areas. Another suggestion is to create multiple instances of the private areas (lounges) in the venue, across the hall would help to break the induced homophily and increase exposure possibilities during journeys between the private areas. Finally, in facilitating for the intimacy gradient, we believe there is a need for a variety of social spaces where elemental encounters could be embedded in a less structured way e.g., coffee area, relaxation area with bean bags, etc.

Understanding Space Syntax Beyond Functionality: Through this work, we attempted to investigate the impact of spatial configuration on the formation of social ties. However, our exploration was limited to only the functional aspect of the place, i.e., we were only informed about the purpose of the different zones and we analysed the collected data through this lens. The functional aspect is one of the many aspects of space syntax. As such, we see opportunities for computational data-driven studies that explore the other dimensions of space syntax (e.g., shape, size, colour, brightness, ambience, architecture, interior etc.) to understand the impact of space phenomena on social behaviour further in different physical settings with different functional purposes.

Limitations

The results and observation presented in this paper should be taken in the light of the type of the event, the venue and the invited participants and thus may not necessarily be generalisable to all events. Furthermore, in this paper we studied the activities of the attendees from a coarse-grain perspective rather than fine-grain interactions amongst individuals. This decision, based on system design consideration, meant that we cannot claim whether a long period of co-location between two individuals corresponds to meaningful interactions between them. However, based on the theories of social psychology we can claim that such proximity does increase the chances of elemental encounter between the pair. Another limitation of our work, is the small population of 61 participants for two special groups with respect to the scale of the events. Although initially we had approval for 100 participants, due to logistical complexities and lack of participation we were confined to a small population of established entrepreneurs and investors depriving us from studying those younger and less established firms in seeking opportunities.

Despite these limitations, we believe our work is the first of its kind to study the impact of space syntax on the entrepreneurs opportunities at a scale of an industrial event with no burden on the participants and by purely relying on ubiquitous Wi-Fi analytics. We expect that our results and recommendation will help future event organisers to successfully design and arrange of networking events of varying scale. We also hope that the system, theoretical framework and insights developed in this work will bring clarity and guidance to aid future UbiComp practitioners seeking to study spatio-social behaviour at a large scale.

REFERENCES

1. Suresh Bhagavatula, Tom Elfring, Aad van Tilburg, and Gerhard G. van de Bunt. 2010. How social and human capital influence opportunity recognition and resource mobilization in India's handloom industry. *Journal of Business Venturing* 25, 3 (2010), 245–260.

2. Ulf Blanke, Tobias Franke, Gerhard Tröster, and Paul Lukowicz. 2014. Capturing crowd dynamics at large scale events using participatory GPS-localization. In *Proceedings of the 9th IEEE International Conference on Intelligent Sensors, Sensor Networks and Information Processing*. 1–7.

3. Peter Michael Blau. 1977. *Inequality and Heterogeneity: A Primitive Theory of Social Structure*. MACMILLAN Company.

4. Stephen P. Borgatti and Daniel S. Halgin. 2011. Analyzing Affiliation Networks. In *The Sage Handbook of Social Network Analysis*, John Scott and Peter J. Carrington (Eds.). SAGE Publications Ltd, London, 417–433.

5. Tina Bratkovic, Bostjan Antoncic, and Mitja Ruzzier. 2009. Strategic utilization of entrepreneur's resource-based social capital and small firm growth. *Journal of Management & Organization* 15, 04 (2009), 486–499.

6. Leo Breiman. 2001. Random Forests. *Machine Learning* 45, 1 (Oct. 2001), 5–32.

7. Chloë Brown, Christos Efstratiou, Ilias Leontiadis, Daniele Quercia, Cecilia Mascolo, James Scott, and Peter Key. 2014. The architecture of innovation: Tracking face-to-face interactions with ubicomp technologies. In *Proceedings of the 2014 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 811–822.

8. Xiang Chen, Yiran Chen, Mian Dong, and Charlie Zhang. 2014. Demystifying Energy Usage in Smartphones. In *Proceedings of the 51st Annual Design Automation Conference*. 1–5.

9. Yohan Chon, Suyeon Kim, Seungwoo Lee, Dongwon Kim, Yungeun Kim, and Hojung Cha. 2014. Sensing WiFi packets in the air: practicality and implications in urban mobility monitoring. In *Proceedings of the International Joint Conference on Pervasive and Ubiquitous Computing*. 189–200.

10. Donna Cox, Volodymyr V. Kindratenko, and David Pointer. 2003. IntelliBadge$^{\mathrm{TM}}$: Towards Providing Location-Aware Value-Added Services at Academic Conferences. In *Porceedings of the 5th International Conference on Ubiquitous Computing*. 264–280.

11. Justin Cranshaw, Eran Toch, Jason Hong, Aniket Kittur, and Norman Sadeh. 2010. Bridging the Gap Between Physical Location and Online Social Networks. In *Proceedings of the 12th ACM International Conference on Ubiquitous Computing*. 119–128.

12. Per Davidsson and Benson Honig. 2003. The role of social and human capital among nascent entrepreneurs. *Journal of Business Venturing* 18, 3 (2003), 301–331.

13. Nathan Eagle and Alex Sandy Pentland. 2009. Eigenbehaviors: Identifying structure in routine. *Behavioral Ecology and Sociobiology* 63, 7 (2009), 1057–1066.

14. Tom Elfring and Willem Hulsink. 2003. Networks in entrepreneurship: The case of high-technology firms. *Small business economics* 21, 4 (2003), 409–422.

15. Biyi Fang, Nicholas D. Lane, Mi Zhang, Aidan Boran, and Fahim Kawsar. 2016. BodyScan: Enabling Radio-based Sensing on Wearable Devices for Contactless Activity and Vital Sign Monitoring. In *Proceedings of The 14th ACM International Conference on Mobile Systems, Applications, and Services*. 97–110.

16. Katherine Faust. 1997. Centrality in affiliation networks. *Social Networks* 19, 2 (1997), 157 – 191.

17. Julien Freudiger. 2015. How Talkative is Your Mobile Device?: An Experimental Study of Wi-Fi Probe Requests. In *Proceedings of the 8th ACM Conference on Security & Privacy in Wireless and Mobile Networks*. 8:1–8:6.

18. Yoav Freund and Robert E. Schapire. 1996. Experiments with a New Boosting Algorithm. In *Proceedings of the 13th International Conference on Machine Learning (ICML 1996)*. 148–156.

19. Roy Friedman, Alex Kogan, and Yevgeny Krivolapov. 2013. On Power and Throughput Tradeoffs of WiFi and Bluetooth in Smartphones. *IEEE Transactions on Mobile Computing* 12, 7 (2013), 1363–1376.

20. Erving Goffman. 1961. *Encounters: Two studies in the sociology of interaction*. Bobbs-Merrill.

21. Jeffrey Hightower and Gaetano Borriello. 2001. Location Systems for Ubiquitous Computing. *Computer* 34, 8 (Aug. 2001), 57–66.

22. Pan Hui, Augustin Chaintreau, James Scott, Richard Gass, Jon Crowcroft, and Christophe Diot. 2005. Pocket Switched Networks and Human Mobility in Conference Environments. In *Proceedings of the 2005 ACM SIGCOMM Workshop on Delay-tolerant Networking*. 244–251.

23. Paul Ingram and Michael W Morris. 2007. Do People Mix at Mixers? Structure, Homophily, and the "Life of the Party". *Administrative Science Quarterly* 52, 4 (2007), 558–585.

24. Shuja Jamil, Anas Basalamah, Ahmed Lbath, and Moustafa Youssef. 2015. Hybrid Participatory Sensing for Analyzing Group Dynamics in the Largest Annual Religious Gathering. In *Proceedings of the 2015 ACM International Joint Conference on Pervasive and Ubiquitous Computing*. 547–558.

25. Vassilis Kostakos, Eamonn O'Neill, Alan Penn, George Roussos, and Dikaios Papadongonas. 2010. Brief Encounters: Sensing, Modeling and Visualizing Urban Mobility and Copresence Networks. *ACM Trans. Computer-Human Interaction* 17, 1, Article 2 (2010), 2:1–2:38 pages.

26. Joseph Lampel and Alan D Meyer. 2008. Field-Configuring Events as Structuring Mechanisms: How Conferences, Ceremonies, and Trade Shows Constitute New Technologies, Industries, and Markets. *Journal of Management Studies* 45, 6 (2008), 1025–1035.

27. Nicholas D Lane, Emiliano Miluzzo, Hong Lu, Daniel Peebles, Tanzeem Choudhury, and Andrew T Campbell. 2010. A survey of mobile phone sensing. *Communications Magazine, IEEE* 48, 9 (2010), 140–150.

28. Jakob Eg Larsen, Piotr Sapiezynski, Arkadiusz Stopczynski, Morten Mørup, and Rasmus Theodorsen. 2013. Crowds, Bluetooth, and Rock'N'Roll: Understanding Music Festival Participant Behavior. In *Proceedings of the 1st ACM International Workshop on Personal Data Meets Distributed Multimedia*. 11–18.

29. Hui Liu, H. Darabi, P. Banerjee, and Jing Liu. 2007. Survey of Wireless Indoor Positioning Techniques and Systems. *IEEE Transactions on Systems, Man and Cybernetics, Part C (Applications and Reviews)* 37, 6 (Nov 2007), 1067–1080.

30. Afra Mashhadi, Akhil Mathur, Marc Van Den Broeck, Geert Vanderhulst, Marc Godon, and Fahim Kawsar. 2016. A Case Study on Capturing and Visualising Face-to-Face Interactions in the Workplace. *Proceeding of the 18th International Conference on Human-Computer Interaction with Mobile Devices and Services* (2016).

31. Miller McPherson and Lynn Smith-Lovin. 1987. Homophily in voluntary organizations: Status distance and the composition of face-to-face groups. *American sociological review* (1987), 370–379.

32. Miller McPherson, Lynn Smith-Lovin, and James M Cook. 2001. Birds of a feather: Homophily in social networks. *Annual review of sociology* (2001), 415–444.

33. Raul Montoliu and Daniel Gatica-Perez. 2010. Discovering Human Places of Interest from Multimodal Mobile Phone Data. In *Proceedings of the 9th International Conference on Mobile and Ubiquitous Multimedia*. 1–10.

34. A. B. M. Musa and Jakob Eriksson. 2012. Tracking Unmodified Smartphones Using Wi-fi Monitors. In *Proceedings of the 10th Conference on Embedded Network Sensor Systems*. 281–294.

35. Janine Nahapiet and Sumantra Ghoshal. 1998. Social capital, intellectual capital, and the organizational advantage. *Academy of management review* 23, 2 (1998), 242–266.

36. Eamonn O'Neill, Vassilis Kostakos, Tim Kindberg, Ava Fatah gen. Schieck, Alan Penn, Danae Stanton Fraser, and Tim Jones. 2006. Instrumenting the City: Developing Methods for Observing and Understanding the Digital Cityscape. In *Porceedings of the 8th International Conference on Ubiquitous Computing*. 315–332.

37. Daniel Roggen, Martin Wirz, Gerhard Tröster, and Dirk Helbing. 2011. Recognition of crowd behavior from mobile sensors with pattern analysis and graph clustering methods. *Networks and Heterogeneous Media* 6, 3 (2011), 521–544.

38. Bernhard Schölkopf and Alexander J. Smola. 2001. *Learning with Kernels: Support Vector Machines, Regularization, Optimization, and Beyond*. MIT Press.

39. Wouter Stam. 2010. Industry event participation and network brokerage among entrepreneurial ventures. *Journal of Management Studies* 47, 4 (2010), 625–653.

40. Wouter Stam, Souren Arzlanian, and Tom Elfring. 2014. Social capital of entrepreneurs and small firm performance: A meta-analysis of contextual and methodological moderators. *Journal of Business Venturing* 29, 1 (2014), 152–173.

41. Hendrik Stange, Thomas Liebig, Dirk Hecker, Gennady Andrienko, and Natalia Andrienko. 2011. Analytical Workflow of Monitoring Human Mobility in Big Event Settings Using Bluetooth. In *Proceedings of the 3rd ACM SIGSPATIAL International Workshop on Indoor Spatial Awareness*. 51–58.

42. Geert Vanderhulst, Afra Mashhadi, Marzieh Dashti, and Fahim Kawsar. 2015. Detecting human encounters from WiFi radio signals. In *Proceedings of the 14th International Conference on Mobile and Ubiquitous Multimedia*. 97–108.

43. Mathias Versichele, Tijs Neutens, Matthias Delafontaine, and Nico Van de Weghe. 2012. The use of Bluetooth for analysing spatiotemporal dynamics of human movement at mass events: a case study of the Ghent Festivities. *Applied Geography* 32, 2 (2012), 208–220.

44. Duncan J Watts. 1999. *Small worlds: the dynamics of networks between order and randomness*. Princeton university press.

45. Jens Weppner and Paul Lukowicz. 2013. Bluetooth based collaborative crowd density estimation with mobile phones. In *Pervasive Computing and Communications (PerCom), 2013 IEEE International Conference on*. 193–200.

46. Giovanni Zanca, Francesco Zorzi, Andrea Zanella, and Michele Zorzi. 2008. Experimental comparison of RSSI-based localization algorithms for indoor wireless sensor networks. In *Proceedings of the workshop on real-world wireless sensor networks*. 1–5.

47. Catherine Zimmer. 1986. Entrepreneurship Through Social Networks. *The art and science of entrepreneurship. Ballinger, Cambridge, MA* (1986), 3–23.