The link stream of contacts in a whole hospital

Lucie Martinet¹,6 · Christophe Crespelle¹,2 · Eric Fleury¹ · Pierre-Yves Boëlle³ · Didier Guillemot⁴,5

Received: 15 May 2018 / Revised: 21 August 2018 / Accepted: 4 September 2018 / Published online: 17 September 2018
© Springer-Verlag GmbH Austria, part of Springer Nature 2018

Abstract
We analyse a huge and very precise trace of contact data collected by a network of sensors during 6 months on the entire population of a rehabilitation hospital. We investigate both the topological structure of the average daily link stream of contacts in the hospital and the temporal structure of the evolution of these contacts hour by hour. Our main aims are to unveil striking properties of these two structures in the considered hospital, and to present a methodology that can be used for analysing any link stream where nodes are classified into groups.

Keywords Link stream · Hospital · Close proximity interaction

1 Introduction

The prevalence of AntiMicrobial Resistant Bacteria (AMRB) has been rising worldwide during the past decades and the resistance rates of major nosocomial pathogens have increased up to alarming levels, implying adverse outcomes for affected patients, such as delays or failures of therapies, prolonged hospitalization stay and increased mortality. Upon colonization by an AMR bacteria, a patient becomes an occult carrier. He is then a potential colonization source for other patients and may also disseminate AMRB into the community while transferred to other facilities. In this context, rehabilitation centres are considered to be a large reservoir of AMRB, offering a great potential for development and dissemination into the community.

The MOSAR project aims at examining the factors determining the dynamics of AMRB spread within healthcare facilities (Obadia et al. 2015a, b). These factors are numerous and complex, but it is widely believed that one support for transmissions of AMRB is close proximity interactions (which we simply call contacts throughout this article) (Halleran 2006; Edmunds et al. 2006; Mossong et al. 2008; Mikolajczyk et al. 2008; Bernard et al. 2009; Gundlapalli et al. 2009; Temime et al. 2009; Polgreen et al. 2010). Then, to further reduce transmission, in addition to classical prevention measures (Muto et al. 2003) (such as admission controls, isolation of carriers and hand hygiene), controlling the flux of interactions within the hospital is considered as the next step (Davis et al. 2004; Wernitz et al. 2005; Nijssen et al. 2005). Indeed, contacts and their dynamics strongly influence how transmission occurs (Pastor-Satorras and Vespignani 2001; Moody 2002; Read and Keeling 2003;
Eames 2008; Smieszek et al. 2009; Salathé and Jones 2010; Stehle et al. 2011a). Yet, contacts are difficult to measure efficiently in practice (Beutels et al. 2006; Read et al. 2012), and they may even be harder to change. Recently, however, advances in communication technologies have made it possible to record close proximity interactions with unprecedented detail, allowing an in-depth view of the structure of contacts in real-life settings (Hui et al. 2005; Eagle and Pentland 2006; Cattuto et al. 2010; Lucet et al. 2012) including in environments critical for spreading of diseases (Salathé et al. 2010, 2011b; Isella et al. 2011; Hornbeck et al. 2012; Vanhems et al. 2013). If such contacts actually support transmission, it may open the way to further improvement in hospital hygiene policy.

In this article, we analyse the contact trace collected on the entire population of a rehabilitation hospital during 6 months between June and November 2009, within the MOSAR project. We focus on a period of 59 days (a bit more than 8 weeks) of the measurement, from July 6th to September 2nd involving 492 individuals, 253 patients and 239 staffs. We describe the methodology we used to uncover the key characteristics of this link stream of contacts and the main results we obtained.

1.1 Our contribution

We analyse separately the graph structure of the average daily link stream of contacts (without taking into account its evolution over time) and the temporal evolution of contacts in the hospital hour by hour. For the first goal, we point out significant differences in the contact profiles of services, as well as in contact patterns of patients and staffs, and we reveal a very special structure of interconnections between the services of the hospital and between the socio-professional categories. Finally, we show that the temporal evolution of the contacts in the hospital presents a clear circadian and weekly pattern, and we unfold the very different behaviours of patients and staffs in this temporal pattern.

1.2 Related works

There have been several recent works using sensor devices to unfold contact patterns among individuals (both graph structure and temporal structure) in environments involving patients or children, which present critical risks for spreading of diseases. The measurement analysed in Stehle et al. (2011b) was made on an entire primary school during 3 days. The experiments described in Isella et al. (2011) and Hornbeck et al. (2012) were both conducted during 1 week in some paediatric ward and the one of Vanhems et al. (2013) took place in a geriatric ward, a kind of service we also have in our study, during 3 days. Finally, for sake of completeness, let us mention that a similar experiment was recently conducted on part of the population of an office building during 2 weeks (Génois et al. 2015). Compared to those works, our analyses present two important advantages. First, the measurement we use was made on a much longer period of time (6 months), which allows to observe weekly pattern and to assess the generality of the conclusions we derive on shorter period of times (like 1 day or 1 week). Second, our measurement is not limited to a specific part of the hospital, it involves all patients and all staffs1 of all services of the hospital, which is a key point to have an accurate view of the actual possibility of spreading into the whole hospital, and even inside a given service. Indeed, these possibilities also depend on the contacts occurring outside the service under study.

2 Preliminaries

The contact data we analyse here were recorded using wireless sensor devices carried by all the participants. Each device a sends one beacon signal containing the ID of a every 30 s and constantly listens for the signals of other devices.2 As a consequence, the signal sent by device a is received by all the other devices b that are close enough to a (typically 1–1.5 m). When this occurs, b records the ID of a together with the timestamps of the reception of the signal. On the technical side, it is worth to note that this short transmission range is obtained by sending low-power signals, which also implies that only signals between sensors that are not separated by a physical obstacle, such as a wall, a door or even a human body, are recorded.

Afterwards, time is sliced in slots of 30 s and we keep, for each slot, the list of pairs \(a, b\) of sensors such that one (a or b) recorded the signal of the other. These pairs are undirected as we do not keep track of whether a received the signal of b or b received the signal of a or both: all these three situations give rise to one single occurrence of the (undirected) pair \(a, b\) in the considered 30 s time slot.

Finally, if a pair \(a, b\) occurs in several consecutive time slots of 30 s, we group all its consecutive occurrences into one single interval of contact. Consequently, the contact data analysed in this article consists of a set of quadruplets \((a, b, t_i, t_f)\), which is called a link stream and is denoted as \(L\) in the following. Note despite the fact that they can be seen as equivalent objects (which justifies that we use the same name for them), the formal definition of link stream we use here differs from the one given in Latapy et al. (2017). The

1 More than 99% of individuals accepted to participate and registered for the experiment.
2 The sending time of the different devices are not synchronised but their internal clocks are.
we have $-\) length, and its non-null length, namely called an interval of contact, or simply a contact, and has a $L$ is $T$ is why in the present context, each quadruplet in study, there are two adjacency pairs, namely $\{a, b\}$ and $\{b, c\}$ (nodes $a$ and $c$ are never in contact between $t_1$ and $t_2$), three contacts (two between nodes $a, b$ and one between nodes $b, c$) and the cumulated length of these contacts on period $T$ is $210$ s (60 s for each of the two contacts between nodes $a, b$ and 90s for the contact between nodes $b, c$). Then, in the full-uniform network formed on period $T$, the values associated with each of the three couples of nodes are 2/3 for the number of adjacency pairs, $3/3 = 1$ for the number of contacts and $210/3 = 70$ s for the cumulated length of contacts

Fig. 1 Sample contact data and full-uniform network. a Sample contact data between three nodes $a, b, c$. b the full-uniform network formed from this contact data on period $T = [t_1, t_2]$. The schema shows seven contacts between nodes $a, b, c$, e.g., the first contact between $a$ and $b$ lasts $90$ s from time $t_1$ to time $t_2$. On the period $T$ of study, there are two adjacency pairs, namely $\{a, b\}$ and $\{b, c\}$ (nodes $a$ and $c$ are never in contact between $t_1$ and $t_2$), three contacts (two between nodes $a, b$ and one between nodes $b, c$) and the cumulated length of these contacts on period $T$ is $210$ s (60 s for each of the two contacts between nodes $a, b$ and 90s for the contact between nodes $b, c$). Then, in the full-uniform network formed on period $T$, the values associated with each of the three couples of nodes are 2/3 for the number of adjacency pairs, $3/3 = 1$ for the number of contacts and $210/3 = 70$ s for the cumulated length of contacts

reason for this is that here, we put the emphasis on the notion of an interval of contact which suits better the nature of our data and the analyses we derive on it, and which is absent in the formalism of Latapy et al. (2017).

Definition 1 A link stream $L = \{(a, b, t_s, t_e)\}$ is a set of quadruplets, where $a$ and $b$ are nodes and $t_s$ and $t_e$ are times, that satisfies the two following conditions:

- for all $(a, b, t_s, t_e) \in L$, we have $t_s < t_e$, and
- if $(a, b, t_s, t_e) \in L$ and $(a, b, t'_s, t'_e) \in L$, then $[t_s, t_e] \cap [t'_s, t'_e] = \emptyset$.

In our context, for each quadruplet $(a, b, t_s, t_e) \in L$, $t_s$ is the starting time of one $30$ s time slot and $t_e$ is the ending time of one $30$ s time slot (see example in Fig. 1). The meaning of quadruplet $(a, b, t_s, t_e)$ is that nodes $a$ and $b$ are in contact during all the time slots between $t_s$ and $t_e$ and that they are not in contact in the $30$ s time slot immediately preceding $t_s$ as well as in the $30$ s time slot immediately following $t_e$. This is why in the present context, each quadruplet in $L$ is called an interval of contact, or simply a contact, and has a non-null length, namely $t_e - t_s$, which is a multiple\(^3\) of $30$ s. The author can refer to Obadia et al. (2015b) for a more detailed description of how contact data was gathered.

Notation 1 (Adjacency pair and length of a contact) For a contact $C = (a, b, t_s, t_e)$, we denote $\text{pair}(C) = \{a, b\}$ the pair of nodes involved in contact $C$, which we call the adjacency pair of $C$, and $\text{length}(C) = t_e - t_s$ the length of contact $C$.

Throughout the article, we often analyse link streams restricted to a specified time period (typically $1$ day or $1$ h). We decline our analyses using three parameters.

Definition 2 For a link stream $L$, we define the three following parameters:

1. Number of adjacency pairs, $\#\text{pairs}(L) = |\text{pairs}(L)|$,
2. Number of contacts, $\#\text{cont}(L) = |L|$ and
3. Cumulated length of contacts, $\text{cumul}\_\text{length}(L) = \sum_{C \in L} \text{length}(C)$.

In the following, we usually use these parameters for link streams obtained by restriction of a larger link stream to a time period (see example in Fig. 1). Moreover, we often use the above notions restricted to the contacts of a designated group of individuals or to the contacts between one single pair of nodes. Moreover, in the rest of the article, we use the notions of semi-contact and adjacency semi-pair instead of contact and adjacency pair. One contact (resp. one adjacency pair) between $a$ and $b$ gives rise to two semi-contacts (resp. two adjacency semi-pairs): one attached to $a$ and one attached to $b$. For sake of vocabulary simplicity, in the following, we use the terms contact and pair instead of semi-contact and semi-pair, but all statistics are actually made using semi-contacts and semi-pairs. The reason for this is that it gives a straightforward meaning to mean statistics per individual.

To analyse the graph structure of contacts within the hospital, we use the aggregated view of a link stream defined below.

Definition 3 (Aggregated network of a link stream) The aggregated network $G$ of a link stream $L$ is the graph $G = (V, E)$, where $V$ is the set of nodes involved in $L$ and $E$ is the set of adjacency pairs of $L$ (see Notation 1). Moreover, in the aggregated network $G$, each adjacency pair $\{u, v\}$ is given two weights denoted $\#\text{cont}_{[u, v]}$ and $\text{cumul}\_\text{length}_{[u, v]}$ and defined as $\#\text{cont}_{[u,v]} = |\{C \in L \mid \text{pair}(C) = \{u, v\}\}|$ and $\text{cumul}\_\text{length}_{[u,v]} = \sum_{C \in L \text{ and } \text{pair}(C) = \{u,v\}} \text{length}(C)$.

Along this article, for sake of comparison, we make extensive use of a uniformised version of the aggregated network of the hospital, which we call the full-uniform network and which is defined as follows (see example in Fig. 1).

Definition 4 (Full-uniform network) The full-uniform network associated to a link stream $L$ is the complete graph on

\(^3\) Note that, because of the way we slice the time in slots of $30$ s, the condition “$t_1 - t_2$ is a multiple of $30$ s” holds for any $t_1$ and for any $t_2$ being the bound of some interval, even if they do not bound the same interval of contact.
vertex set \( V \), with \( V \) being the set of nodes involved in \( L \), where each pair of nodes \( u, v \in V \) receives three quantities, which have the same value for all pairs of nodes of \( V \):

1. A fractional number \( \#\text{adj}(u, v) \) of adjacency pairs, between 0 and 1, equal to the density of adjacency pairs between nodes of \( V \), i.e., \( \#\text{adj}(u, v) = 2 \#\text{pairs}(L)/(|V|(|V| - 1)) \),
2. A number of contacts \( \#\text{cont}(u, v) \) equal to the mean number of contacts per pair of nodes, i.e., \( \#\text{cont}(u, v) = 2 \#\text{cont}(L)/(|V|(|V| - 1)) \),
3. A cumulated length of contact \( \text{cumul}_\text{length}(u, v) \) equal to the mean cumulated length per pair of nodes, i.e., \( \text{cumul}_\text{length}(u, v) = 2 \text{cumul}_\text{length}(L)/(|V|(|V| - 1)) \).

Note that we do not use the full-uniform network only to compare the value attached to one pair to the mean value in the hospital. We instead use it in a richer way, as a point of comparison for all the statistics we compute, within one group, between two groups, between one group and the rest of the hospital, etc. This allows us to reveal how specific are the configurations encountered in the hospital, by quantifying how much they deviate from the average behaviour. Along the article, we do the same using other kinds of partially uniformised networks as a point of comparison, where some properties are fixed as in the real network and some others are uniformised. This allows to target more specifically the impact of each property on the structure observed in the real network.

### 2.1 General organisation of the hospital

Over the period of study, the mean number of people present in the hospital during 1 day is about 103 patients and 64 staffs. The patients and staffs are divided into nine services. The average daily size of each service is given in Fig. 2 and their functions are given in Table 1. Only the first five of them (S1–S5) are hospital wards and contain both patients and staffs, the other four (S6–S9) contain only staffs. Each of the five services involving patients occupy one floor in one of the two wings (called Ménard and Sorrel) of the building (see map in Fig. 3). The Sorrel wing hosts three services: nutritional readaptation (S1, 1st floor), neuro-orthopedic reeducation (S2, 2nd floor), and geriatric readaptation (S3, 3rd floor). The Ménard wing hosts the chronic vegetative state service (S4, 2nd floor) and the neurologic readaptation service (S5, 3rd floor). The four services containing only staffs are the night service (S6), regrouping people replacing staffs from services S1 to S5 during nights, and the reeducation services (S7–S9): physiotherapeutic service (S7), ergotherapeutic service (S8) and other reeducation staff service (S9). S7 and S8 are located in two distinct places between the two wings of the buildings, but S6 and S9 do not have a determined location in the hospital. It must be clear that the division of the hospital into services is not meaningful only from an administrative and management point of view but also clearly impacts the structure of the contacts: in average in 1 day, 66% of the adjacency pairs of the hospital occur inside services, 88% for the number of contacts and 92% for the cumulated length of contacts, while these values are only 25% in the full-uniform network.
3 Different levels of activity of services

Figure 4 shows the separation and division of contacts among the nine services of the hospital, in terms of number of adjacency pairs (a), number of contacts (b) and cumulated length of contacts (c). It reveals some important differences between services. The five services including patients seem to be more active than the four others, for each of the three criteria. But clear differences also appear between these five services. As one may guess, one reason for this is that services have different sizes (see Fig. 2). For adjacency pairs, this is confirmed by the fact that the number of mean adjacency pairs per individual per day varies only a little between two different services (Fig. 5a). On the other hand, the number of contacts and the cumulated length of contacts per day remain very different from one service to another even when computed in average for one individual (Fig. 5b, c). This indicates that for these two criteria, the sizes of services cannot be held entirely responsible for the disparities in the activity of services appearing on Fig. 4.

Services S6–S9, which do not include any patient, have a mean number of contacts and a mean cumulated length of contacts per individual which is far smaller than those of services S1–S5, which do include patients (Fig. 5b, c). Moreover, among these latter services, it appears that services S4, S5 and S2 present a higher mean individual activity, for these two parameters, than services S1 and S3; and it turns out that S4, S5 and S2 are the three services that contain the greater number of patients (see Fig. 2).

These observations suggest that the individual activity of patients with regard to number of contacts and cumulated length of contacts may be much higher than the one of staffs.

Another interesting fact revealed by Figs. 4 and 5 is that the number of contacts and the cumulated length of contacts per service behave very similarly. This is confirmed by Fig. 6a which shows the scatter plot of the mean values per individual and per day of these two parameters, for each service of the hospital. The plot shows that for all services, the mean values of number of contacts and cumulated length of contacts are strongly correlated. This is actually an even more general fact as the correlation between these two parameters is also clearly visible for each individual on the whole period of study (Fig. 6b). Therefore, as they give very similar results in all the experiments we conducted, we choose to keep only one of them in most of the analysis presented in the rest of the paper, namely the cumulated length of contacts.

4 Different behaviours of patients and staffs

As pointed out above, patients and staffs seem to have a very different activities. We then refine our analysis of the mean activity per individual and per day by separating patients from staffs in the five concerned services (Fig. 7). It turns out that patients are a bit less active than staffs of the same service (about 20–30% less) in terms of adjacency pairs, but are much more active in terms of cumulated length of contact (between two and six times more, except for service
S1 where cumulated length of contacts of patients and staffs are comparable). This explains why the differences between services that appeared on Fig. 4a for the whole service disappear when considering the adjacency pairs per person (Fig. 5a), while this difference does not disappear for cumulated length (see Figs. 4c, 5c). Indeed, as the number of adjacency pairs is comparable for patients and staffs, so are the mean individual values for each service, independently of the fact that they contain more patients or more staffs. On the opposite, for cumulated length of contacts, as the one of patients is much higher than the one of staffs, it follows that the mean individual value in services containing both staffs and patients (S2–S5, S1 being an exception as pointed above) is higher than the mean individual value in services containing only staffs (S6–S9).

These differences rise some crucial questions: what is the role of patients and staffs in the global contact pattern of the hospital? Where is the majority of contacts located? between patients, between staffs or between patients and staffs? Table 2 shows that a vast majority of the cumulated length of contacts in the hospital, 80%, involves two patients, while only 12% of this length involve one patient and one staff, and 8% involve two staffs. Nevertheless, the picture for adjacency pairs is quite different: those between patients represent only 24% of all pairs, which is about 35% less than in the full-uniform network. The majority of adjacency pairs, 56%, involves one patient and one staff, and 8% involve two staffs. Nevertheless, the picture for adjacency pairs is quite different: those between patients represent only 24% of all pairs, which is about 35% less than in the full-uniform network. The majority of adjacency pairs, 56%, involves one patient and one staff, and 20% of them involve two staffs. Both of these values are about 20% higher than in the full-uniform network. This suggests that the contacts of staffs and in particular the contacts between staffs and patients are very important for the structure of the daily contact network, and may then play a key role regarding the possibility of spreading in the hospital.

Table 2 Distribution of contacts between patients and staffs in the hospital

|        | PA-PA | PA-ST | ST-ST |
|--------|-------|-------|-------|
| (a) Global distribution |        |       |       |
| Pairs  | 0.24  | 0.56  | 0.20  |
| Length | 0.80  | 0.12  | 0.08  |
| PA vs PA |       |       |       |
| PA vs ST |       |       |       |
| (b) Patient centred |        |       |       |
| Pairs  | 0.46  | 0.54  |       |
| Length | 0.93  | 0.07  |       |
| ST vs PA |       |       |       |
| ST vs ST |       |       |       |
| (c) Staff centred |        |       |       |
| Pairs  | 0.58  | 0.42  |       |
| Length | 0.42  | 0.58  |       |

Fig. 6 Scatter plots of the number of contacts vs the cumulated length of contacts.
(a) Mean values per individual and per day for each service, b values for each individual in the hospital on the whole period of experiment

Fig. 7 Mean activity per individual and per day for each service.
(a) Number of adjacency pairs, b cumulated length of contacts. Patients are in light pink and staffs in dark turquoise

Table 2 Distribution of contacts between patients and staffs in the hospital
Tables 2b, c give the separation and division of contacts, respectively, for an average patient and an average staff. They show that the majority of the adjacency pairs of a patient (54%) occurs with a staff, and that the majority of the adjacency pairs of a staff (58%) occurs with a patient. Note that, opposite to the case of patients whose cumulated length of contacts is strongly unbalanced in favour of contacts with patients (93%), staffs share much more equitably their length of contacts between patients (42%) and staffs (58%). This confirms that staffs present a more open pattern of contacts than the one of patients, which may result for them in particular spreading abilities.

5 Introversion and interconnection of services

We mentioned previously that most of the activity of the link stream takes place inside services. Here we investigate further this question by examining the introversion of each service with regard to adjacency pairs, number of contacts and cumulated length of contacts. In the rest of the article, we qualify contacts and adjacency pairs as internal or external depending on whether they take place inside a service or between two distinct services. Figure 8 gives the separation and division for each service between its internal and external activity, in terms of adjacency pairs (a), number of contacts (b) and cumulated length of contacts (c). One can clearly distinguish two groups of services on the three plots: services S1–S5 which have the bigger part of their activity occurring inside the service itself and services S6 to S9 which have the bigger part of their activity occurring outside the service. For number of contacts (b) and cumulated length of contacts (c), for all services, the proportion of internal activity is augmented compared to the proportion of internal adjacency pairs. For services S1–S5, all of these values are above 80%, while they are between 18 and 53% only for services S6–S9. These differences are partly explained by the fact that services S6–S7 contains only staffs, including many healthcare workers, but no patients. It is then natural that a large proportion of the contacts of the members of these services occurs outside the service, for example, with individuals from the services containing patients. But there are other factors that may explain the differences observed between services S1–S5 and services S6–S9 in the separation and division between their internal and external activity. The size of the service, for example, has a strong impact on this division, as smaller services are likely to have a more important part of their activity directed outside of the service. One way to separate the contribution of the size of services in our analysis, and then isolate the contribution of the functional characteristics of services, is to introduce the notion of factor of introversion.

Formally, for $a$ being one of the three parameters we use (namely adjacency pairs, number of contacts and cumulated length of contacts), the int/ext ratio of a service $S$ with regard to $a$ is defined as $\alpha_{\text{int}}(S)/\alpha_{\text{ext}}(S)$, where $\alpha_{\text{int}}(S)$ is the value of parameter $a$ (e.g., number of adjacency pairs) inside $S$ and $\alpha_{\text{ext}}(S)$ is the value of parameter $a$ between $S$ and the rest of the hospital. Then, the factor of deviation of the int/ext ratio of service $S$, which we also call factor of introversion, is defined as the quotient between the int/ext ratio of $S$ in the real aggregated network and the int/ext ratio of $S$ in some specifically defined uniform network. For adjacency pairs, we use for comparison the full-uniform network (cf. Definition 4). For number of contacts, we use the contact-uniform network (cf. Definition 5 below), which has exactly the same adjacency pairs as the real aggregated network, and for cumulated length of contacts we use the length-uniform network (cf. Definition 6 below), which has the same adjacency pairs as the real aggregated network, each of which has the same number of contacts as in the real aggregated network.

Definition 5 (Contact-uniform network) The contact-uniform network associated to a link stream $L$ is the graph
$G = (V, E)$, with $V$ being the set of nodes involved in $L$ and $E$ being the set of adjacency pairs of $L$, where each adjacency pair $\{u, v\}$ is assigned the same number of contacts $\#\text{cont}(u,v) = \#\text{cont}(L)/\#\text{pairs}(L)$, which is the mean number of contacts per adjacency pairs in $L$.

**Definition 6** (Length-uniform network) The length-uniform network associated to a link stream $L$ is the graph $G = (V, E)$, with $V$ being the set of nodes involved in $L$ and $E$ being the set of adjacency pairs of $L$, where each adjacency pair $\{u, v\}$ is assigned its number $\#\text{cont}_{(u,v)}$ of contacts in $L$ and a cumulated length of contact $\text{cumul\_length}(u,v) = \#\text{cont}_{(u,v)} \cdot (\text{cumul\_length}(L)/\#\text{cont}(L))$, which is the cumulated length obtained for $\{u, v\}$ by making all contact lengths equal to the mean length of contacts of $L$.

The rationale behind these definitions is that for the number of contacts, we compute its deviation knowing the adjacency pairs of the real aggregated network, and for the cumulated length of contacts, we compute its deviation knowing both the adjacency pairs and the number of contacts of the real aggregated network. This allows to determine to which extent the deviation observed for one parameter (e.g., number of contacts) is a consequence of the deviation observed for other parameters (e.g., adjacency pairs), or is independent from it.

The results are depicted on Fig. 9. They confirm, as observed on Fig. 8, that services S1–S5 are strongly introverted in terms of adjacency pairs: they have a factor of introversion between 9 and 18. More surprisingly, services S7 and S8 also appear to be strongly introverted, which was not predictable from Fig. 8. Their factor of introversion are both above 15, meaning that the ratio of adjacency pairs between inside and outside the service is 15 times higher (in favour of internal pairs) that what it would be if contacts between individuals in the hospital occurred completely freely, independently of spatial, organisational and functional constraints. Even services S6 and S9, which do not have a single determined location in the building of the hospital, also appear to be more introverted than expected in the full-uniform network, their factor of introversion being higher than 2. Then, despite of what we could expect from Fig. 8, all services strongly favour adjacency pairs inside the service rather than outside, in a very unbalanced way for at least 7 out of 9 of them which have a factor of deviation higher than 9.

Going further from (Fig. 9b), even knowing this unbalanced structure of the adjacency pairs, services are still clearly introverted in terms of number of contacts (factors between 1.5 and 5). This means that services do not have only a strong preference for making adjacency pairs inside rather than outside, but they are also much more likely to repeat contacts for their internal adjacency pairs. For cumulated length (Fig. 9c), the factor of introversion is less than 2 for all services, but always strictly greater than 1. The fact that these values are lower than the previous ones is a consequence of the correlation between cumulated length of contacts and number of contacts (see Fig. 6). But still, they indicate that services not only favour internal adjacency pairs and internal repetition of contacts, but also prefer longer contacts between their members rather than outside.

Table 3 gives some global statistics distinguishing both between internal and external adjacency pairs and between patients and staffs. It reveals a strong bipartite-like structure of the aggregated network between the staffs divided into services on one side (nine classes), and the patients divided into services on the other side (five classes), see Fig. 10. Indeed, more than 83% ($=0.56/(0.56 + 0.05 + 0.06)$) of the adjacency pairs between these 14 classes occur between one patient and one staff. In addition, links between patients and staffs represent more than 67% of the external links between services of the hospital (18% of these links occur between staffs and 15% between patients). This shows that the contacts between patients and staffs play a prevalent role in connecting the introverted services of the hospital. These observations are confirmed from an individual-centred point of view (see Table 3b, c): an individual (either patient or staff) has only few external adjacency pairs with his own side of the bipartition, while the distribution between its external and internal pairs with the other side are more balanced than internal/external pairs in the whole hospital.

Table 4 gives the same kind of statistics as Table 3, but considering cumulated length of contacts instead of adjacency pairs. The resulting picture of the hospital is quite different. First, 78% of the total cumulated length of contacts in the hospital occurs between two patients in a same service,
which was not at all the case for adjacency pairs. We give more explanation about this fact in Sect. 7, dedicated to the temporal structure of contacts, by considering the times of the day when these contacts occur. Second, patients also play a more important role in the cumulated length of contacts between two different services: 87% of this length is made by contacts involving at least one patient and this proportion is shared in a balanced way between contacts involving two patients and contacts involving one patient and one staff. The importance of cumulated length of contacts involving patients for connecting different services is confirmed by the patient centred view (Table 4b) and the staff centred view (Table 4c). Omitting the time one patient spends with patients of his/her service, the rest of his cumulated length of contacts is very equitably shared between staffs of his/her own service, staffs of other services and patients of other services. A staff spends a large proportion of its cumulated length of contacts with staffs of his service (50%) and only few with staffs of other services (8%). Staffs also clearly favour cumulated length of contacts with patients inside their service rather than outside. This shows, that opposite to the case of adjacency pairs, where pairs involving staffs seem particularly important in connecting the different services between them, for cumulated length of contacts, it is the contacts of patients that seem to be prevalent between different services. This probably results in qualitatively different abilities of patients and staffs for spreading diffusions in the hospital.

6 Affinities between services

The question we address in this section is to determine whether some pairs of services are more likely to interact between them than others. Figure 11a gives the average density of adjacency pairs per day between every pair of

| PA-PA | PA-ST | ST-ST | All |
|-------|-------|-------|-----|
| Ext.  | 0.05  | 0.23  | 0.06| 0.34|
| Int.  | 0.19  | 0.33  | 0.14| 0.66|
| All   | 0.24  | 0.56  | 0.20| 1.00|

| PA vs PA | PA ST | ST All |
|----------|-------|--------|
| Ext.     | 0.10  | 0.22   | 0.32 |
| Int.     | 0.36  | 0.32   | 0.68 |
| All      | 0.46  | 0.54   | 1    |

| ST vs PA | PA ST | ST All |
|----------|-------|--------|
| Ext.     | 0.24  | 0.12   | 0.36 |
| Int.     | 0.34  | 0.30   | 0.64 |
| All      | 0.58  | 0.42   | 1    |

Fig. 10 Bipartite-like structure of the aggregated network between classes of patients and classes of staffs. One class is constituted either by the patients (up) or by the staffs (bottom) of one single service. Black bold lines are for links crossing the bipartition, grey lines for

Table 3 Distribution of adjacency pairs between patients and staffs, distinguishing between internal and external pairs

|         | PA-PA   | PA-ST | ST-ST | All   |
|---------|---------|-------|-------|-------|
| Ext.    | 0.05    | 0.23  | 0.06  | 0.34  |
| Int.    | 0.19    | 0.33  | 0.14  | 0.66  |
| All     | 0.24    | 0.56  | 0.20  | 1.00  |
services $S_i S_j$, i.e., the number of adjacency pairs in 1 day between $S_i$ and $S_j$ divided by the number of possible pairs on this day $|S_i||S_j|$. Similarly, Fig. 12a gives the average cumulated length of contacts per day and per pair of individuals between $S_i$ and $S_j$. The rationale for dividing by the number of possible pairs between services $S_i$ and $S_j$ is that the values obtained describe the intensity of interactions between service $S_i$ and $S_j$, independently of their size. As we have shown earlier, all services do not have the same activity level. Naturally, two services $S_i$ and $S_j$ having higher activity levels will tend to have a higher intensity of interactions between them. This does not mean that $S_i$ and $S_j$ favour the interactions between them compared to interactions with other services: this higher intensity simply results from the fact that both of them interact more with everyone in the hospital. With the aim of constraining or rearranging contacts inside the hospital for limiting the possible diffusions of bacteriological strains, it is highly desirable to be able to identify which pairs of services favour the interactions between them compared to their interactions with other services: this higher intensity simply results from the fact that both of them interact more with everyone in the hospital. With the aim of constraining or rearranging contacts inside the hospital for limiting the possible diffusions of bacteriological strains, it is highly desirable to be able to identify which pairs of services favour the interactions between them compared to their interactions with other services: this higher intensity simply results from the fact that both of them interact more with everyone in the hospital. With the aim of constraining or rearranging contacts inside the hospital for limiting the possible diffusions of bacteriological strains, it is highly desirable to be able to identify which pairs of services favour the interactions between them compared to their interactions with other services: this higher intensity simply results from the fact that both of them interact more with everyone in the hospital. With the aim of constraining or rearranging contacts inside the hospital for limiting the possible diffusions of bacteriological strains, it is highly desirable to be able to identify which pairs of services favour the interactions between them compared to their interactions with other services: this higher intensity simply results from the fact that both of them interact more with everyone in the hospital. With the aim of constraining or rearranging contacts inside the hospital for limiting the possible diffusions of bacteriological strains, it is highly desirable to be able to identify which pairs of services favour the interactions between them compared to their interactions with other services: this higher intensity simply results from the fact that both of them interact more with everyone in the hospital. With the aim of constraining or rearranging contacts inside the hospital for limiting the possible diffusions of bacteriological strains, it is highly desirable to be able to identify which pairs of services favour the interactions between them compared to their interactions with other services: this higher intensity simply results from the fact that both of them interact more with everyone in the hospital. With the aim of constraining or rearranging contacts inside the hospital for limiting the possible diffusions of bacteriological strains, it is highly desirable to be able to identify which pairs of services favour the interactions between them compared to their interactions with other services: this higher intensity simply results from the fact that both of them interact more with everyone in the hospital. With the aim of constraining or rearranging contacts inside the hospital for limiting the possible diffusions of bacteriological strains, it is highly desirable to be able to identify which pairs of services favour the interactions between them compared to their interactions with other services: this higher intensity simply results from the fact that both of them interact more with everyone in the hospital. With the aim of constraining or rearranging contacts inside the hospital for limiting the possible diffusions of bacteriological strains, it is highly desirable to be able to identify which pairs of services favour the interactions between them compared to their interactions with other services: this higher intensity simply results from the fact that both of them interact more with everyone in the hospital. With the aim of constraining or rearranging contacts inside the hospital for limiting the possible diffusions of bacteriological strains, it is highly desirable to be able to identify which pairs of services favour the interactions between them compared to their interactions with other services: this higher intensity simply results from the fact that both of them interact more with everyone in the hospital. With the aim of constraining or rearranging contacts inside the hospital for limiting the possible diffusions of bacteriological strains, it is highly desirable to be able to identify which pairs of services favour the interactions between them compared to their interactions with other services: this higher intensity simply results from the fact that both of them interact more with everyone in the hospital. With the aim of constraining or rearranging contacts inside the hospital for limiting the possible diffusions of bacteriological strains, it is highly desirable to be able to identify which pairs of services favour the interactions between them compared to their interactions with other services: this higher intensity simply results from the fact that both of them interact more with everyone in the hospital. With the aim of constraining or rearranging contacts inside the hospital for limiting the possible diffusions of bacteriological strains, it is highly desirable to be able to identify which pairs of services favour the interactions between them compared to their interactions with other services: this higher intensity simply results from the fact that both of them interact more with everyone in the hospital. With the aim of constraining or rearranging contacts inside the hospital for limiting the possible diffusions of bacteriological strains, it is highly desirable to be able to identify which pairs of services favour the interactions between them compared to their interactions with other services: this higher intensity simply results from the fact that both of them interact more with everyone in the hospital. With the aim of constraining or rearranging contacts inside the hospital for limiting the possible diffusions of bacteriological strains, it is highly desirable to be able to identify which pairs of services favour the interactions between them compared to their interactions with other services: this higher intensity simply results from the fact that both of them interact more with everyone in the hospital. With the aim of constraining or rearranging contacts inside the hospital for limiting the possible diffusions of bacteriological strains, it is highly desirable to be able to identify which pairs of services favour the interactions between them compared to their interactions with other services: this higher intensity simply results from the fact that both of them interact more with everyone in the hospital. With the aim of constraining or rearranging contacts inside the hospital for limiting the possible diffusions of bacteriological strains, it is highly desirable to be able to identify which pairs of services favour the interactions between them compared to their interactions with other services: this higher intensity simply results from the fact that both of them interact more with everyone in the hospital. With the aim of constraining or rearranging contacts inside the hospital for limiting the possible diffusions of bacteriological strains, it is highly desirable to be able to identify which pairs of services favour the interactions between them compared to their interactions with other services: this higher intensity simply results from the fact that both of them interact more with everyone in the hospital. With the aim of constraining or rearranging contacts inside the hospital for limiting the possible diffusions of bacteriological strains, it is highly desirable to be able to identify which pairs of services favour the interactions between them compared to their interactions with other services: this higher intensity simply results from the fact that both of them interact more with everyone in the hospital. With the aim of constraining or rearranging contacts inside the hospital for limiting the possible diffusions of bacteriological strains, it is highly desirable to be able to identify which pairs of services favour the interactions between them compared to their interactions with other services: this higher intensity simply results from the fact that both of them interact more with everyone in the hospital. With the aim of constraining or rearranging contacts inside the hospital for limiting the possible diffusions of bacteriological strains, it is highly desirable to be able to identify which pairs of services favour the interactions between them compared to their interactions with other services: this higher intensity simply results from the fact that both of them interact more with everyone in the hospital. With the aim of constraining or rearranging contacts inside the hospital for limiting the possible diffusions of bacteriological strains, it is highly desirable to be able to identify which pairs of services favour the interactions between them compared to their interactions with other services: this higher intensity simply results from the fact that both of them interact more with everyone in the hospital. With the aim of constraining or rearranging contacts inside the hospital for limiting the possible diffusions of bacteriological strains, it is highly desirable to be able to identify which pairs of services favour the interactions between them compared to their interactions with other services: this higher intensity simply results from the fact that both of them interact more with everyone in the hospital. With the aim of constraining or rearra...
contacts. If we denote $|D_i|$ the number of semi-pairs assigned to $S_i$, in average, the expected number of pairs between $S_i$ and $S_j$ resulting from this random matching process is $|D_i||D_j|/|D|$, where $|D|$ is the total number of semi-pairs in the whole hospital. One can proceed similarly with duration of contacts instead of adjacency pairs. In this case, the expected cumulated length of contacts between $S_i$ and $S_j$ is $|E_i||E_j|/|E|$, denoting $|E_i|$ the cumulated length of semi-contacts of $S_i$ and $|E|$ the total cumulated length of all semi-contacts in the hospital. This average network is precisely the one we use for computing deviations. The meaning of the deviation factor is then to show which pairs of services have more (deviation factor greater than 1) or less (deviation factor less than 1) interactions between them that what would be expected from their respective activity levels if they shared their activity uniformly with other services (that is without tendency to favour or unfavour contacts with some services). When the deviation factor between $S_i$ and $S_j$ is greater than 1, we say that the relationship between $S_i$ and $S_j$ is favoured and when their deviation factor is less than 1 we say that their relationship is unfavoured.

Looking at the statistics for adjacency pairs (Fig. 11a), it appears that the large majority of adjacency pairs occurs between individuals of services S1–S5 on one side, which are the services containing both patients and staffs, and individuals of services S6–S9 on the other side, which contain only staffs. Looking closer, and taking into account the deviations of these values (Fig. 11b), one can actually distinguish finer groups of services. The first two groups are S1, S2, S3 and S4, S5. While the interactions inside each of these groups are high, it turns out that the deviation factors between the two groups have high unfavoured values, meaning that services S1, S2, S3 have a strong tendency not to interact with S4 and S5. This probably comes from the fact that each of these two groups of services occupies a distinct wing of the building, which are physically separated (see map in Fig. 3). The third group contains only service S6 which interacts a lot with services S1–S5 but has almost no interactions with services S7–S9. The reason is that S6 is the night service and therefore involves staffs working in other services of the hospital, except the reeducation services (S7–S9) which work only at daytime. Finally, services S7–S9, strongly connected together, constitute the fourth group.

The picture given by the cumulated length of contacts is similar, with some meaningful differences. Note that the deviations for cumulated length of contacts are often higher than the one observed for adjacency pairs: the favoured and unfavoured relationships are more clearly marked when seen through duration of contacts. On the other hand, the general organisation into groups highlighted above through adjacency pairs is still visible on cumulated length of contacts.

The main difference is that more than half of the relationships between services S1–S5 and services S6–S7 appear to be unfavoured in terms of cumulated length of contacts while they were almost all favoured in terms of adjacency pairs, revealing a more complex pattern of contacts in the hospital.

To get a deeper insight into the structure of the relationships between services of the hospital, we used both adjacency pairs and cumulated length of contacts to build the graph of favoured relationships between services on Fig. 13a. In the drawing, only pairs of services $S_iS_j$ that have a clearly favoured relationship are linked by a line, i.e., both deviation factors, for adjacency pairs and cumulated length, between $S_i$ and $S_j$ are greater than 1 and at least one of them is greater than 1.5. Moreover, if both deviation factors are greater than 1.5, then the line linking $S_i$ and $S_j$ on the drawing is thicker and darker and we say that the relationship is strongly favoured. One can retrieve on Fig. 13a the details of the big picture of relationships between services described above. Similarly, Fig. 13b shows the graph of unfavoured relationships between services. It is remarkable that the relationships between services of the hospital are strongly

![Fig. 12 Affinities between services based on cumulated length of contacts.](image-url)
polarised: 29 pairs of services out of 36 possible pairs are either clearly favoured (15 of them, and 8 of them are even strongly favoured) or clearly unfavoured (14 of them, and 11 of them are strongly unfavoured). Moreover, looking closer, only 3 out of the 7 remaining pairs (i.e., about 8% of the total number of possible pairs) have a balanced deviation on the favoured and on the unfavoured sides (for adjacency pairs on one side and for cumulated length of contacts on the other side) that forbids to mark them as favoured or unfavoured. The four others are strongly deviated on one side and only slightly deviated on the other side. These pairs may then safely be classified as favoured (such as S7S1) or as unfavoured (such as S7S2, S9S2, S9S5). Totally, more than 90% of the relationships between services of the hospital are polarised (and more than 80% are clearly polarised in the sense above). This shows that taking into account the structure of relationships between services in the analysis of the difusions occurring in the hospital is certainly relevant.

As we pointed out earlier, services are widely introverted and the connections between them strongly relies on contacts between staffs and patients. To get a deeper insight into the structure of interconnections of services through contacts between patients and staffs, we apply the methodology above restricted to these contacts. Fig. 14 gives the intensity of relationships and their deviation factors for adjacency pairs and Fig. 15 for the cumulated length of contacts.

We retrieve the big picture revealed above (see Figs. 11, 12) with some interesting differences. First, the relationships between services S1 and S5 are in general more clearly marked, both favourably or unfavourably, on contacts between patients and staffs. It is also remarkable that almost all these relationships are symmetric (except the one between S2 and S3): if patients of S_i have a clearly favoured relationship with staffs of S_j, then so have the staffs of S_i with the patients of S_j. Nevertheless, these clearly favoured relationships may be quite unbalanced (as for S1S2 and S1S3): patients of S_i favour contacts with staffs of S_j much more than staffs of S_i do with patients of S_j.
For services S6–S9, the main difference observed is for the night service S6: most of its clearly favoured relationships with services S2–S5 disappear when considering only contacts with patients of S2–S5. This shows that an important part of the contacts of the night service occurs with staffs of services S1–S5 and fewer with patients. The pattern of contacts between reeducation services (S7–S9) and healthcare services is also slightly changed when considering only patient–staff contacts. Moreover, as noted for the interactions between the services S1 and S5, the strength of their favoured or unfavoured relationships (see Figs. 14b, 15b) are higher for patient–staff interactions than what they were at the level of the whole services. This confirms our previous observation on the importance of contacts between patients and staffs and shows that these contacts strongly shape the structure of contacts in the hospital.

7 Organisation of the hospital with regard to socio-professional categories

In this section, we investigate the structure of contacts between socio-professional categories of staffs. To this purpose, we apply to the staffs divided into socio-professional categories the same methodology we used for the whole hospital divided into services. The specific organisation of the contacts of staffs is of high interest for epidemiological issues. First, the structure of these contacts is deeply constrained by the role of each category of staffs and this strongly shapes the general picture of all the contacts in the hospital. Second, contacts of staffs are more likely to be changed by changing the organisation and policy of the hospital, while only limited constraints can be applied to patients.

In the hospital of Berck-sur-mer, staffs are divided into 12 different professional categories. Their titles are given in Table 5 and their sizes, in average number of individuals per day, are given on Fig. 16. It shows that most of the categories have very few representatives a day (at most four) except categories C1–C3, which concentrate a large majority of the staffs of the hospital. We can a priori distinguish three groups of socio-professional categories which have similar functions in the hospital. The first group is made of categories C1–C5, which are nurses and housekeepers. These categories are the more present and visible in the daily life of the hospital and potentially constantly working in contact with other people in the hospital, especially patients. The second functional group is the one made of categories C7–C9 which contain reeducation staffs. These categories are specialised and give more occasional care to patients. Finally, categories C11 and C12 are logistic and administrative staffs of the hospital. They are less mobile and less likely to be in contact with patients. Categories C6 (stretcher bearers) and C10 (organisers and hairdressers) have roles that do not fit directly into the functional groups given above. Nevertheless, their type of activity, devoted to specific and occasional services to patients, makes them closer to the group of reeducation staffs. The order we chose for categories, which is the same in all the figures below, respects the functional groups identified above and is at the same time adapted to the groups identified by our subsequent analyses.

Figure 17 shows the average daily activity of one staff in each category, in terms of number of adjacency pairs (a) and cumulated length of contacts (b). It takes into account both the contacts between two staffs and the contacts between one staff and one patient. Both plots reveal two groups of categories having different levels of activity. In Fig. 17a, categories C1–C9 appear to have a higher and rather similar number of adjacency pairs, while this number is relatively lower for categories C10–C12. For cumulated length of contacts, the two groups are slightly different than those for adjacency pairs, and the difference between them is much more clearly marked: staffs in categories C1–C7 have a much longer cumulated length of contacts than staffs in categories C8–C12. Staffs of C11 and C12, logistic and administration, can even be distinguished from the rest of this group by their particularly short daily cumulated length of contacts. In summary, both adjacency pairs and cumulated length of contacts reveal that staffs whose function is devoted to daily care or rehabilitation of patients have a higher level of activity than other categories, therefore implying a different exposition to diffusion for socio-professional categories.
As done for services in Sect. 5, Fig. 18a gives the average density of adjacency pairs per day between every pair of categories $C_iC_j$, i.e., the number of adjacency pairs in 1 day between $C_i$ and $C_j$ divided by the number of possible pairs on this day $|C_i||C_j|$. Figure 18b gives the deviation of these values compared to the configuration model. Figure 19a, b give the same for cumulated length of contacts. As usual, the density shows where the interactions are more intense, and its deviation shows which interactions are favoured, i.e., the affinities between categories.

The big pictures arising from the observation of these statistics for adjacency pairs and for cumulated length of contacts are essentially the same, though it appears more clearly for cumulated length of contacts. They reveal a quite specific structure of contacts between socio-professional categories as the contacts between staffs are not at all equitably shared between categories. The most intense relationships as well as the most favoured relationships actually occur within three (overlapping) groups of categories: the first group formed by categories C1–C5, the second group by categories C5–C8, and the third one by categories C9–C12. This partition of the categories in three groups is very clearly marked. It is emphasized by the fact that categories C1–C3 have strongly unfavoured relationships with categories C6–C8 (see Figs. 18b, 19b) and by the fact that categories C11 and C12 have unfavoured relationships with most of the other categories (see Fig. 19b), except C9. This group structure is articulated by C5, which belongs to both the first and second group (with a preference for the second one), and by the couple C8, C9, which have a strong affinity in terms of cumulated length of contacts (see Fig. 19b) and which then create a bridge between the second and the third group. The particular position of C9 is even accentuated by the fact that it is strongly tied with the rest of the third group. Outbound of this partition into three groups, C9 and C10 also plays a transversal role by having some affinities with categories belonging to other groups, both for adjacency pairs and for cumulated length of contacts. This structure therefore confers key roles to categories C5, C8, C9 and C10, whose impact on diffusion properties of the link stream of contacts in the whole hospital is worth investigating. Finally, another interesting fact revealed by these analyses is that categories of the first group, C1–C5, are the only ones that unfavour contacts within their own category, both for adjacency pairs and for cumulated length of contacts. The impact of this on the diffusions occurring into this group should also be deeper investigated.

8 Temporal structure of contacts in the hospital

In this section we study the evolution over time of contacts in the hospital. Figure 20 gives the evolution over 1 week (from Monday July 6th to Sunday July 12th) of the number of active individuals per hour, i.e., the individuals who had at least one contact during the considered hour. One retrieves the circadian rhythm of many human activities: the number of active individuals is higher between 6:00 and 24:00, with a peak around 12:00, when people gather for lunch, and this number is much lower between 0:00 and 6:00. There is also a clear weekly pattern denoted by a lower number of active individuals during weekends. Figure 20 also distinguishes between the number of active patients and the number of active staffs. Both of these numbers observe the same circadian and weekly pattern. Nevertheless, the variation of the number of active staffs is much higher than the variation of the number of active patients. As one may guess, the main reason for this is that patients stay in the hospital all day (and night) long, while there are much less staffs in the hospital at night (only staffs of the night service, S6). Moreover, this phenomenon is strengthened by a specific pattern of contacts for patients that we point out below.

In Fig. 21, we plot the evolution, during the same week, of the number of adjacency pairs per hour, number of contacts per hour and cumulated length of contacts per hour, in the whole hospital. All these three parameters follow the pattern highlighted above on the number of active individuals, denoting a very strong temporal structure of the activity of the hospital. Nevertheless, it is worth noticing that this
pattern appears more clearly for the number of adjacency pairs and the number of contacts. The reason is that the gap between the cumulated length of contacts during daytime and during night time is not as wide as for the two other parameters. In particular, the cumulated length of contacts has high values both in the morning, between 6:00 and 12:00, and in the evening between 18:00 and 24:00. We give more explanations of this fact in the following by considering separately the activity of patients and the activity of staffs. Let us emphasize the fact that the results we obtain here are not particular to the week we consider and hold for all the weeks of the period of study.

Figure 22 shows the number of adjacency pairs per individual per hour (degree), cumulated length of contacts per individual per hour and the mean cumulated length of contact for an adjacency pair, separating patients (a) from staffs (b). Remember that only individuals involved in at least one contact during the considered hour are taken into account (see Definition 3). The reason for this is that here, we focus on the contact pattern of the active part of the link stream. It appears that both plots follow the circadian and weekly pattern of the whole hospital. As we pointed out previously (cf. Fig. 7), the average cumulated length of contacts for patients is much higher than the one of staffs. On the other hand, for staffs, this cumulated length varies more in time than the one of patients: there are several peaks of activity for staffs in 1 day while there is mainly one for patients, moreover, the value for staffs regularly becomes very low (less than 5 min per hour) while the value for patients almost always remain above 10 min per hour. Looking closer at the times when the peaks of activity occur, one can see that, surprisingly, the cumulated length of contacts of active patients is higher
Fig. 19  Affinities between categories based on cumulated length of contacts. a Mean cumulated length of contacts (in s) per pair of staffs and per day between pairs of categories, b factors of deviation of these values. We use exponent −1 to denote the inverse of a number, e.g., 2.6−1 means 1/2.6, and ∞ stands for an infinite value. The blue scale of colours is for favoured relationships and the red one for unfavoured relationships. Both matrices are symmetric.

| C1  | C2  | C3  | C4  | C5  | C6  | C7  | C8  | C9  | C10 | C11 | C12 |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 54  | 120 | 120 | 148 | 77  | 5.9 | 1.7 | 0.9 | 16  | 67  | 3.2 | 54  |
| 120 | 52  | 112 | 168 | 66  | 8.7 | 3.1 | 3.2 | 19  | 67  | 3.2 | 54  |
| 120 | 112 | 71  | 89  | 133 | 5.7 | 1.2 | 0.4 | 20  | 42  | 3.0 | 1.1 |
| 148 | 168 | 89  | 171 | 162 | 98  | 164 | 9.0 | 10  | 97  | 2.2 | 0.0 |
| 77  | 66  | 133 | 162 | 114 | 225 | 412 | 95  | 38  | 20  | 10  | 6.3 |
| 5.9 | 8.7 | 5.7 | 98  | 225 | 454 | 1143| 40  | 10  | 244 | 19  | 0.0 |
| 1.7 | 3.1 | 1.2 | 164 | 412 | 1143| 514 | 52  | 26  | 0.0  | 1.3 | 0.0 |
| 0.9 | 3.2 | 0.4 | 9.0 | 95  | 40  | 52  | 606 | 56  | 0.0  | 7.4 | 1.2 |
| 15  | 19  | 20  | 10  | 38  | 10  | 26  | 56  | 78  | 26  | 1.1 | 228 |
| 34  | 67  | 42  | 97  | 20  | 244 | 0.0  | 0.0  | 26  | 0.0  | 3.5 | 0.0 |
| 2.1 | 3.2 | 3.0 | 2.2 | 10  | 19  | 1.3 | 7.4 | 7.1  | 0.0  | 170 | 19  |
| 1.3 | 54  | 1.1 | 0.0 | 6.9 | 0.0 | 0.0  | 1.2 | 228 | 3.5  | 19  | 335 |

Fig. 20  Number of active individuals per hour in the hospital. Patients are in light pink, staffs in dark turquoise and the sum of all individuals is depicted by a thin black line.
during night (between 22:00 and 8:00) and much lower during days (between 10:00 and 20:00). For active staffs, the situation is opposite. Their cumulated length of contacts is very low at night (between 0:00 and 6:00) and their peaks of activity generally occur around 8:00, 12:00, 20:00 and 23:00. This observation has to be tempered with the fact that only active patients are taken into account and that their number is lower during nights, see Fig. 20.

Nevertheless, interestingly, this difference can be further explained considering the curves of number of pairs and mean cumulated length per pair. For patients, the cumulated length of contacts varies opposite to the number of adjacency pairs, but follows the mean cumulated length per pair. During day time, patients are in contact with several persons and only very few at night time. But at night, the mean cumulated length of their adjacency pairs is longer. The reason for this is that most of them share their room and therefore have very long contacts with their roommate at night, as the distance between beds is usually no more than 1.5 m. For staffs the situation is very different. Their mean cumulated length of contact per adjacency pair appears to be much more stable along time than the one of patients. Moreover, the variations of their cumulated length of contacts are not opposed to the variations of their number of adjacency pairs (as it is the case for patients) but are rather in accordance with them, despite the fact that they also show some visible differences.

The activity pattern of staffs appears clearly on Fig. 23. It shows, separately for active patients (a) and active staffs (b), the variations of the cumulated length of contacts compared with those of the number of contacts and mean duration per contact. One can notice that, opposite to the cumulated length of contacts, the number of contacts is higher for staffs than for patients. On Fig. 23b, it is striking to see that the cumulated length of contacts of staffs very closely follows their number of contacts. In the meanwhile, the mean duration of their contacts appears quite stable along time with an average value lower than the one of patients (see Fig. 23a). This shows that for staffs, the cumulated length of their contacts is made by the repetition of numerous contacts (with many different persons, see Fig. 22 and the discussion above) and not by longer contacts with a few number of persons, as it is the case for patients. Figure 23a confirms this fact: the cumulated length of contacts of patients does not follow at all their number of contacts, but is in accordance with the mean duration of their contacts.

Thus, the daily pattern of contacts for patients and staffs is drastically different. For patients, their cumulated length of contacts depends on the cumulated length of their adjacency pairs: it is high during nights when they have long contacts with a very restricted number of persons (usually only one or two). For staffs, their mean cumulated length of contacts per adjacency pair as well as the mean duration of their contacts do not vary much: they have longer time of
contacts when they have more numerous contacts, which happens several times per day at rather fixed times around 8:00, 12:00, 20:00 and 23:00. These deep differences are very likely to have a strong impact on the way patients and staffs can propagate spreadings in the hospital.

9 Applications

The ultimate aim of the analyses conducted in this article is to understand and control the spread of nosocomial infections in hospitals. In the context of the MOSAR project, we wish more specifically to understand how Staphylococcus aureus can become resistant to methicillin. To that purpose, building models of dissemination in the hospital is a very important step to be able to compare different scenarios and understand better some of the factors that have an impact on the dissemination process. To obtain efficient and representative models, one needs to carefully identify and quantify the factors to be introduced. By establishing a very precise map of the contacts between the services of the hospital and by revealing some of their key characteristics, our work is a decisive step toward the conception of accurate models of dissemination. Note that, to fully achieve this goal, the knowledge about the structure of contacts must be coupled with the factors relative to the behaviour of bacteria, that have to be determined besides.

Another very concrete application of our work is that it can be used to adjust the protection policy implemented in the hospital, with the aim of limiting further the possibilities of spreading. Indeed, in the picture of the preferred interactions between services that we draw, some are difficult to constrain because they are related to the care given in the hospital, while some others are induced by the administrative or spatial organisation of the hospital. Part of these latter contacts may then be avoided by a change in this organisation. For the contacts related to care, which cannot be avoided, their particular situation in the global structure of contacts in the hospital make them more or less critical for conveying spreadings in the whole structure. The contacts that are more critical could therefore be applied to an increased level of hygiene precautions compared to the standard procedures. The map of contacts we draw at the level of the hospital gives practitioners crucial information to be able to decide to implement such relevant protection policies for careworkers and patients.

Finally, let us note that the type of approach we follow here is relevant to be applied in various other contexts where contacts are the main vector of some diffusion process, such as the study of the dynamic learning process of language at school or the diffusion of opinions and information in an organisation, for instance.

10 Generality of our study and limitations

It must be clear that the results we presented here are specific to the hospital of Berck-sur-Mer where the measurement was done. The structure of contacts in another hospital would very probably show different characteristics, since our study pointed out that this structure is strongly marked by the spatial and administrative organisation of the hospital. There are also some other conclusions that may very likely be general to other environments, such as the differences in the behaviour of patients and staff, and the different patterns of activity depending on the time of the day. The size of the hospital may also play a role in shaping the contacts between individuals. Therefore, conducting similar studies in environments of different sizes would be of interest.

The characteristics of the measurement method may have an impact on the conclusions obtained here as well. First, the short period of sampling, set here at 30 s, still misses some very quick contacts, happening at a lower time scale. It would be desirable to collect and analyse data with a finer measurement period to assess whether these short contacts affect the global structure pictured here. Another key characteristic of the measurement is the range of detection of sensors. It is believed that the most relevant contacts for the spread of nosocomial infections are the very close interactions. In the perspective of testing this hypothesis, it would be very interesting to be able to shorten even more the range of detections of the sensors used in the measurement and to observe the impact of it on the structure of contacts observed in the whole hospital.

11 Conclusion and perspectives

We presented here the first analysis of the link stream of contacts in a whole hospital during a long-term period. We designed a method to investigate both the temporal and topological structures of this link stream. Our method constitutes a diagnostic tool that can be used to reveal the main characteristics of the contacts with regard to the organisation of the hospital within services and socio-professional categories. It is generic and can be applied to any hospital, and more generally to any link stream where nodes are a priori partitioned into functional groups.

The application of this method to the link stream of contacts in the hospital of Berck-sur-Mer allows to objectively describe and precisely quantify the structure of contacts in the hospital. In particular, we are able to draw a map of the preferred interactions between the different services. As a result of these analyses, we also provide some important observations for understanding and controlling the spread of nosocomial infections in hospitals. First, contacts are not
at all uniformly shared between services and between socio-professional categories. The analyses we conducted following these two dimensions show a very specific and strongly marked structure of the contacts in the hospital. This pattern of contacts, which may be different for each considered hospital but which also probably exists for each of them, certainly plays a key role in the propagation of infections within the hospital. Another observation, is that this pattern is different depending on whether one considers daily adjacency pairs or cumulated length of contacts. This points out that it is crucial to clarify the impact of duration of contacts on the possibility of transmission, as depending on the importance of this impact, the possibilities of diffusions within the hospital may appear quite different. Our analyses confirm that patients and staffs exhibit a quite different pattern of contacts. But contrary to what was observed in some earlier measurements (Vanhems et al. 2013), in terms of cumulated length, the contacts between patients play a very important role for connecting the whole hospital, both inside services and between services.

Moreover, from a temporal point of view as well, the structure of the contacts in the hospital appears complex and clearly marked. It exhibits a strong heterogeneity which certainly has a great impact on diffusions within the hospital. Then, any policy aiming at limiting such diffusions should imperatively take into account the temporal dimension of contacts within the hospital. In addition, we note that the activity during nights should not be neglected: the major part of the duration of contacts of patients occurs during this time. The impact of this long duration proximity at night between patients sharing the same room must be further investigated for airborne diseases.

Clearly, the main perspective of our work is to determine the impact of the specific structure of contacts we highlighted on spreading processes. The challenge lies in the fact that the dataset contains only the contaminated nodes and their time of contamination, as it is impossible, with current technology, to measure through which contacts the contamination occurred. Therefore, a first step toward this goal is to use synthetic diffusion processes on the link stream of contacts to determine what are the contacts more likely to propagate diffusions, and then check whether contaminated nodes of the dataset are more often involved in such contacts than non-contaminated nodes are.

Acknowledgements The authors thank all the I-Bird (Individual-Based Investigation of Resistance Dissemination) study group members.

Funding This study was supported by the European Commission under the Life Science Health Priority of the 6th Framework Program (MOSAR network contract LSHP-CT-2007-037941). This work was performed within the framework of the LABEX MILYON (ANR-10-LABX-0070) of Université de Lyon, within the program “Investissements d’Avenir” (ANR-11-IDEX-0007) operated by the French National Research Agency (ANR). This work was performed within the framework of the LABEX IBEID (ANR-10-LABX-62). The second author gratefully acknowledges the support from a Grant from Région Rhône-Alpes and from the delegation program of CNRS.

References

Bernard H, Fischer R, Mikolajczyk RT, Kretzschmar M, Wildner M (2009) Nurses’ contacts and potential for infectious disease transmission. Emerg Infect Dis 15(9):1438–1444

Beutels P, Shkedy Z, Aerts M, Van Damme P (2006) Social mixing patterns for transmission models of close contact infections: exploring self-evaluation and diary-based data collection through a web-based interface. Epidemiol Infect 134:1158–1166

Cattuto C, Van den Broeck W, Barrat A, Colizza V, Pinton JF, Vespignani A (2010) Dynamics of person-to-person interactions from distributed RFID sensor networks. PLoS One 5:e11596

Davis K, Stewart J, Crouch H, Florez C, Hospenthal D (2004) Methicillin-resistant Staphylococcus aureus (MRSA) nares colonization at hospital admission and its effect on subsequent MRSA infection. Clin Infect Dis 39(6):776–782

Eagle N, Pentland A (2006) Reality mining: sensing complex social systems. Person Ubiquitous Comput 10(4):255–268

Eames K (2008) Modelling disease spread through random and regular contacts in clustered populations. Theor Popul Biol 73(1):104–111

Edmunds W, Kafatos G, Wallinga J, Mossong J (2006) Mixing patterns and the spread of close-contact infectious diseases. Emerg Themes Epidemiol 3:10

Génois M, Vestergaard CL, Fournet J, Panisson A, Bonmarin I, Barrat A (2015) Data on face-to-face contacts in an office building suggest a low-cost vaccination strategy based on community linkers. Netw Sci 3(3):326–347

Gundlaptalli A, Ma X, Benuzillo J, Petey W, Greenberg R, Hales J, Leecaster M, Samore M (2009) Social network analyses of patient–healthcare worker interactions: implications for disease transmission. In: AMIA annual symposium proceedings. pp 213–217

Halloran E (2006) Invited commentary: challenges of using contact data to understand acute respiratory disease transmission. Am J Epidemiol 164(10):945–946

Hornbeck T, Naylor D, Segre A, Thomas G, Herman T, Polgreen P (2012) Using sensor networks to study the effect of peripatetic healthcare workers on the spread of hospital-associated infections. J Infect Dis 206:1549–1557

Hui P, Chaintreau A, Scott J, Gass R, Crowcroft J, Diot C (2005) Pocket switched networks and human mobility in conference environments. In: ACM SIGCOMM workshop on delay-tolerant networking (WDTN 2005). ACM, pp 244–251

Iselfa L, Romano M, Barrat A, Cattuto C, Colizza V, Van den Broeck W, Gesualdo F, Pandolfi E, Ravà L, Rizzo C, Tozzi AE (2011) Close encounters in a pediatric ward: measuring face-to-face proximity and mixing patterns with wearable sensors. PLoS One 6(2):e17144

Latapy M, Viard T, Magnien C (2017) Stream graphs and link streams for the modeling of interactions over time. CoRR abs/1710.04073

Lucet JC, Laouenan C, Chelius G, Veziris N, Lepelletier D, Friggeri A, Abiteboul D, Bouvet E, Mentré F, Fleury E (2012) Electronic sensors for assessing interactions between healthcare workers and patients under airborne precautions. PLoS One 7(5):e37893

Martinet L, Crespelle C, fleury E (2014) Dynamic contact network analysis in hospital wards. In: 5th workshop on complex networks (ComNet2014). Springer, no. 549 in studies in computational intelligence. pp 241–249

Springer
Mikolajczyk RT, Akmatov MK, Rastin S, Kretzschmar M (2008) Social contacts of school children and the transmission of respiratory-spread pathogens. Epidemiol Infect 136(6):813–822
Molloy M, Reed B (1995) A critical point for random graphs with a given degree sequence. Rand Struct Algorithms 6(2–3):161–180
Moody J (2002) The importance of relationship timing for diffusion. Soc Forces 81(1):25–56
Mossong J, Hens N, Jit M, Beutels P, Auranen K, Mikolajczyk R, Massari M, Salmaso S, Tomba GS, Wallinga J, Heijne J, Sadkowska-Todys M, Rosinska M, Edmunds WJ (2008) Social contacts and mixing patterns relevant to the spread of infectious diseases. PLoS Med 5(3):e74
Muto CA, Jernigan JA, Ostrowsky BE, Richet HM, Jarvis WR, Boyce JM, Farr BM (2003) SHEA guideline for preventing nosocomial transmission of multidrug-resistant strains of Staphylococcus aureus and enterococcus. Infect Control Hosp Epidemiol 24(5):362–386
Nijssen S, Bonten M, Weinstein R (2005) Are active microbiological surveillance and subsequent isolation needed to prevent the spread of methicillin-resistant Staphylococcus aureus? Clin Infect Dis 40(3):405–409
Obadia T, Opatowski L, Temime L, Herrmann JL, Fleury E, Boelle PY, Guilleminot D (2015a) Interindividual contacts and carriage of methicillin-resistant Staphylococcus aureus: a nested case-control study. Infec Control Hosp Epidemiol 36(8):922–929
Obadia T, Silhol R, Opatowski L, Temime L, Legrand J, Thiébaut A, Herrmann JL, Fleury E, Guilleminot D, Boëlle PY (2015b) Detailed contact data and the dissemination of Staphylococcus aureus in hospitals. PLoS Comput Biol 11(3):1–16
Pastor-Satorras R, Vespignani A (2001) Epidemic spreading in scale-free networks. Phys Rev Lett 86(14):3200–3203
Polgreen PM, Tassier TL, Pemmaraju SV, Segre AM (2010) Prioritizing healthcare worker vaccinations on the basis of social network analysis. Infect Control Hosp Epidemiol 31(9):893–900
Read JM, Keeling MJ (2003) Disease evolution on networks: the role of contact structure. Proc R Soc B Biol Sci 270(1516):699–708
Read JM, Edmunds WJ, Riley S, Lessler J, Cummings DA (2012) Close encounters of the infectious kind: methods to measure social mixing behaviour. Epidemiol Infect 140(12):2117–2130
Salathé M, Jones JH (2010) Dynamics and control of diseases in networks with community structure. PLOS Comput Biol 6(4):e1000736
Salathé M, Kazandjieva M, Lee JW, Levis P, Feldman MW, Jones JH (2010) A high-resolution human contact network for infectious disease transmission. PNAS 107(51):22020–22025
Smieszek T, Fiebig L, Scholz RW (2009) Models of epidemics: when contact repetition and clustering should be included. Theor Biol Med Model 6:11
Stehle J, Voirin N, Barrat A, Cattuto C, Colizza V, Isella L, Regis C, Pinton JF, Khanafer N, Van den Broeck W, Vanhems P (2011a) Simulation of an SEIR infectious disease model on the dynamic contact network of conference attendees. BMC Med 9:87
Stehle J, Voirin N, Barrat A, Cattuto C, Isella L, Pinton JF, Quaggiotto M, den Broeck WV, Régis C, Lina B, Vanhems P (2011b) High-resolution measurements of face-to-face contact patterns in a primary school. PLoS One 6(8):e23176
Temime L, Opatowski L, Pannet Y, Brun-Buisson C, Boëlle PY, Guilleminot D (2009) Peripatetic health-care workers as potential super-spreaders. PNAS 106(43):18420–18425
Vanhems P, Barrat A, Cattuto C, Pinton JF, Khanafer N, Régis C, Ba K, Comte B, Voirin N (2013) Estimating potential infection transmission routes in hospital wards using wearable proximity sensors. PLoS One 8(9):e73970
Wernitz M, Swidsinski S, Weist K, Sohr D, Witte W, Roloff KFD, Ruden H, Veit S (2005) Effectiveness of a hospital-wide selective screening programme for methicillin-resistant Staphylococcus aureus (MRSA) carriers at hospital admission to prevent hospital-acquired MRSA infections. Clin Microbiol Infect 11(6):457–465