Supporting Information

Salathé et al. 10.1073/pnas.1009094108

SI Methods

Data Collection. ***General.*** On January 14, 2010, we distributed wireless senor network motes (TelosB; Crossbow Technologies Inc.) to all students, teachers, and staff at an American high school (the date was chosen because it represented a typical school day). Participants were asked to sign an assent form on which they indicated at what time the mote was turned on. The assent form also asked participants to indicate their role/status at the school, with the following four options available: "student," "teacher," "staff," and "other." At the end of the day, we collected the motes and assent forms and then obtained data with the corresponding assent from 789 motes/individuals. Some of the motes had not been used (because of people either being absent from the school or not participating in the project), and we did not obtain written assent to use the data for some motes with data. The remaining data cover 94% of the entire school population. We also deployed motes at fixed locations (stationary motes), but these are not part of the dataset described here except for one stationary mote in the main cafeteria; the signal of this mote was used to reconstruct the global timestamp (see below).

Deployment details. Motes were distributed in batches (with an average of ∼11 motes) the night before the deployment and handed out to participants starting around 6:00 AM (with the vast majority of participants receiving and activating their motes on arrival at 8:00 AM). Participants were asked to put their mote in a thin plastic pouch attached to a lanyard (provided by us) and to wear the lanyard around the neck, with the mote being located in front of the chest at all times. The participants handed the motes back to us when leaving the school or at the end of the school day (the vast majority was received between 4:00 and 4:30 PM). The technical details regarding code design, signal strength considerations, and other issues have been described elsewhere (1); however, briefly, each participant's mote was programmed to broadcast beacons at −16.9 dBm at a regular 20-s interval; the packet included the sender's local sequence number. On receiving a beacon, the mote checked the received signal strength indicator (RSSI) value of the packet. Note that the motes are always scanning; thus, no interactions with a duration of at least 20 s will be missed. If the signal strength was lower than −80 dBm, the packet was discarded (this decision was based on experimental data showing that when subjects were facing each other, packets within 3 m had an RSSI value of roughly −80 dBm or above; packets sent when one subject was facing the other person's back had a lower RSSI value (1) (Fig. S4). Otherwise, the receiver created a contact entry, consisting of the sender's ID and beacon sequence number as well as the local mote's sequence number and the RSSI value of the packet.

TelosB motes have a 1-MB flash memory in which interactions can be stored, thus eliminating the need to broadcast interactions to any other external hardware for storage. As a consequence, interactions between subjects can be captured anywhere on the campus of the school, an area of more than 45,000 m^2.

Reconstructing the full contact network required a global timestamp, relative to which all interactions between subjects occurred. Local sequence numbers in each data trace acted as relative clocks, and they could be used as offsets from one stationary mote (the "master stationary mote," located in the main cafeteria), which would be the master clock providing global time. Packets originating from this mote were transmitted at high power (−11 dBm) and were not subject to RSSI filtering at the receiver. More than 90% of mobile motes had received one or more beacons from the master stationary mote. For these mobile motes, we calculated the offset between the master and local sequence numbers. In addition, we created a table of offsets to serve as a lookup table, which included mobile motes as well as other stationary motes. To process data traces from mobile motes that did not hear directly from the master stationary mote, we used the offsets table transitively to compute a timestamp from another mote that already had its global time.

After processing the raw data, we thus obtained a list of interactions that contains 762,868 unique interactions between two motes for a duration of n consecutive beacon intervals (Datasets S1, S2, and S3). Because beacons are broadcast every 20 s, the number of beacons can be used as an approximate measure of contact duration (such that duration in minutes was approximately $n/3$).

This project was approved by the Stanford University Institutional Review Board on July 24, 2009.

1. Kazandjieva M, et al. (2010) Experiences in measuring a human contact network for epidemiology research. *HotEmNets '10: Proceedings of the ACM Workshop on Hot Topics in Embedded Networked Sensors*, (Association for Computing Machinery, Killarney, Ireland).

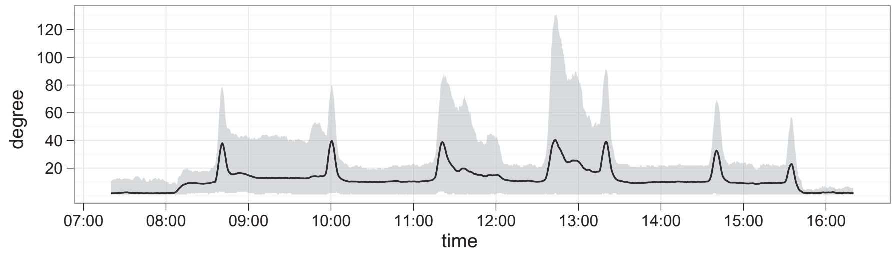

Fig. S1. Temporal dynamics of the average number of contacts (degree). Here, the degree of an individual is measured as the number of other individuals in close proximity during 5 min. Gray background spans the 2.5% and 97.5% percentiles of the degree distribution.

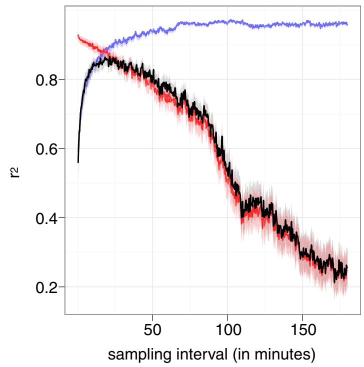

Fig. S2. Squared correlation (r^2) between outbreak size and degree of index case (black), outbreak size and strength of index case (red), and degree and strength of index case (blue) at various sampling rates. The left-most correlations are based on the full dataset (sampling interval of 1/3 min), and all others are based on subsampled datasets that would have been generated with the given sampling interval. The shaded area behind the line shows the 95% confidence interval of squared correlation.

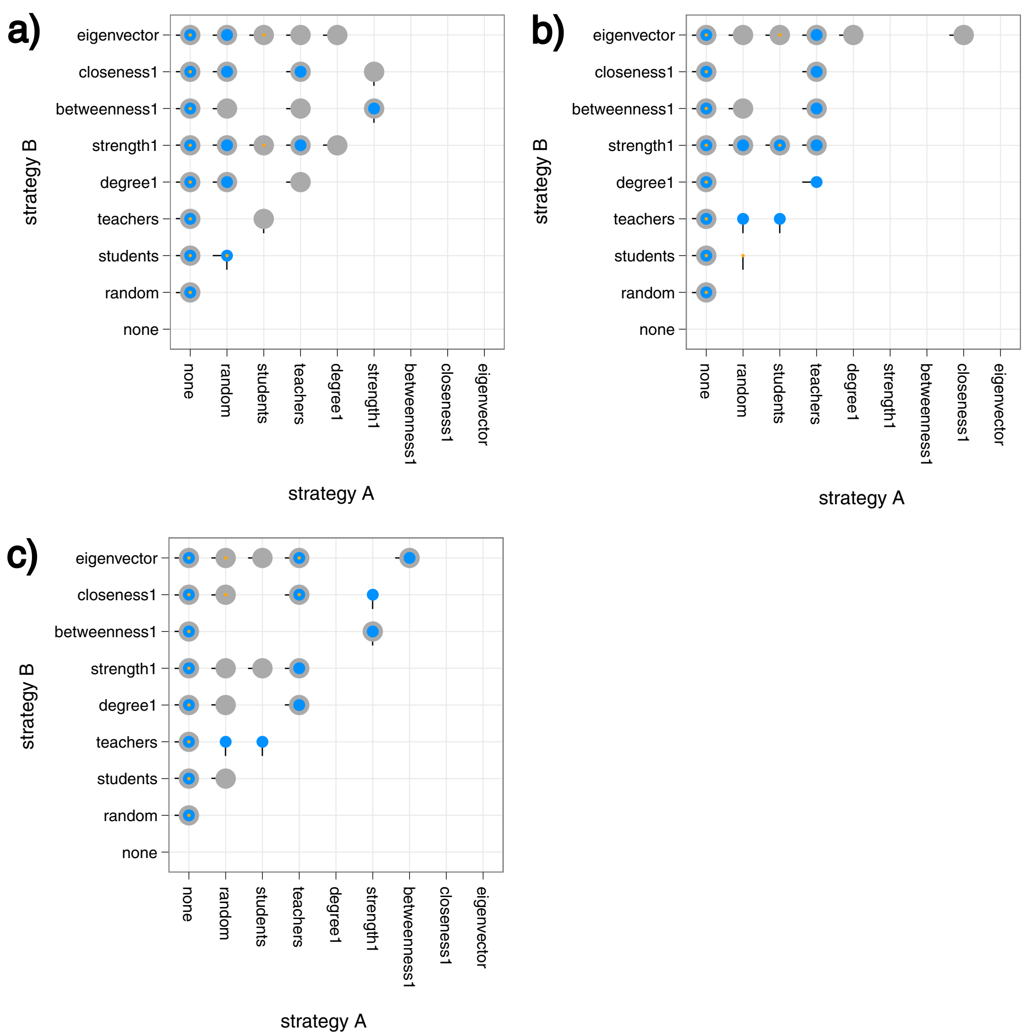

Fig. S3. Settings identical to those described in Fig. 4*B*, but the results are separated according to transmission probabilities per CPR used [0.002 (*A*), 0.003 (*B*), and 0.0045 (*C*)].

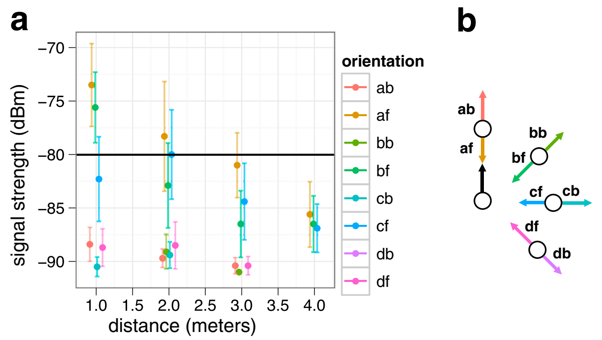

Fig. S4. Dependency of signal strength on distance (1, 2, 3, and 4 m), orientation (a, b, c, and d forward or backward), and angle (0°, 45°, 90°, and 135°). The black horizontal line shows the threshold value that was chosen for the data collection. (*A*) Points show the average signal strength, and bars represent the SD of a particular measurement. Some settings lack data because no packets were received. (A slight horizontal offset was added to the data points for visual clarity.) (*B*) Spatial setting of the four angles and two directions are shown, with reference to the main mote.

Other Supporting Information Files

Dataset S1 (TXT)
Dataset S2 (TXT)
Dataset S3 (TXT)