The eigenmode analysis of human motion

Juyong Park\textsuperscript{1}, Deok-Sun Lee\textsuperscript{2} and Marta C González\textsuperscript{3}

\begin{itemize}
\item \textsuperscript{1} Department of Physics, Kyung Hee University, Seoul 130-701, Republic of Korea
\item \textsuperscript{2} Departments of Natural Medical Sciences and Physics, Inha University, Incheon 402-751, Republic of Korea
\item \textsuperscript{3} Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
\end{itemize}

E-mail: perturbation@gmail.com (J Park), deoksun.lee@inha.ac.kr and martag@mit.edu

Received 10 August 2010
Accepted 11 October 2010
Published 15 November 2010

Online at stacks.iop.org/JSTAT/2010/P11021
doi:10.1088/1742-5468/2010/11/P11021

Abstract. Rapid advances in modern communication technology are enabling the accumulation of large-scale, high-resolution observational data of the spatiotemporal movements of humans. Classification and prediction of human mobility based on the analysis of such data has great potential in applications such as urban planning in addition to being a subject of theoretical interest. A robust theoretical framework is therefore required to study and properly understand human motion. Here we perform the eigenmode analysis of human motion data gathered from mobile communication records, which allows us to explore the scaling properties and characteristics of human motion.

Keywords: network dynamics, random graphs, networks, stochastic processes, traffic and crowd dynamics
1. Introduction

Thanks to rapid advances in data retrieval technology coupled with the advent of the Internet and personal electronic devices, we are witnessing the emergence of unprecedented opportunities to collect, analyze, and understand massive data encoding a variety of human behaviors, prompting calls for collaborative efforts from various disciplines in the scientific community [1]. One topic gaining increased interest in the physics community in particular is how one can understand the spatiotemporal movements of humans [2], an example being a recent numerical study by González et al of human mobility data gathered from mobile phone communication records [3]. One intriguing finding in their work was that the distance or area covered by human trajectories appears to be bounded or increase extremely slowly in time $t$, which highlights a stark difference from classical models of particle motion such as random walk and Lévy Flight [4, 5] where the expected fluctuation in the displacement of a particle grow unbounded as polynomials of time $t$ (a notable exception is the random walk in random environment (RWRE) where a weak spatial randomness assigned to lattice points lead to a $\sim \log^2 t$ behavior in the average fluctuation [6–10]). Given the success of these models in explaining real physical systems such as diffusive gases in free space, it is natural to examine their validity in understanding human mobility data. What González et al showed, in essence, is that the simplistic nature of these models is not suited for that task, demonstrating the inherent complexity of human motion. For the same reason, devising a simple generative model of human motion appears extremely difficult, if at all possible, requiring new theoretical approaches to understanding the observed behavior of human motion on a deeper level.

As a first step in that direction, in this paper we study analytically core observations given in González et al. Specifically, we employ the framework of the Markov Process, a fundamental and indispensable methodology in the study of stochastic processes found in a variety of scientific and engineering fields such as linguistics, computational biology, graph
The eigenmode analysis of human motion

Figure 1. (a) The format of data used in our analysis. For each mobile user, a table lists the time and location, represented by the nearest transmission tower IDs, of \( N \) calls initiated by the individual. \( K \), the number of unique tower IDs, satisfies \( K \leq N \). (b) A graphical representation of the data. Transmission towers distributed over the 2D space constitute a mesh of non-overlapping Voronoi cells.

theory, and information theory, to name but a few [11]–[14]. We construct the Markov transition matrix from mobile communication data, and investigate its properties via eigenmode analysis. We demonstrate that it allows us to understand the aforementioned slow increase of fluctuations in human motion in time and, furthermore, that one can characterize individual motions via the eigenmodes of the transition matrices.

This paper is organized as follows. In section 2, we introduce the Markov transition matrix constructed from the observed transition rates between spatial coordinates. In section 3, we investigate the various properties of its eigenmodes, from which we present a theoretical understanding of the radius of gyration. In section 4, we show that the eigenmodes also act as concise descriptors of the characteristics of individual movement patterns. Finally, in section 5 we summarize our findings and discuss their implications.

2. The human mobility data and the Markov chain framework

For this study, we employed a mobile phone-usage data set provided by a western European service provider, used in a previous study of González et al [3]. Figure 1(a) illustrates the basic format of the data; for each individual a table lists the time and the location of each of \( N \) calls initiated by the user (varying from individual to individual) during an observational period of six months. The spatial location of the user at the call initiation is given by the ID of the transmission tower nearest to the user, meaning that the resolution of the user location is dictated by the size of non-overlapping Voronoi cells around each tower, presented schematically in figure 1(b). Since one Voronoi cell corresponds to one
transmission tower, we may use the two concepts interchangeably. The number \( K \) of unique tower IDs appearing in figure 1(a) satisfies \( K \leq N \), the number of calls. From this table, we build for each individual a \( K \times K \) Markov matrix (also called the transition matrix) \( \mathbf{M} = \{m_{ij} \equiv m_{i \leftarrow j} | 1 \leq i, j \leq K \} \), where \( m_{ij} \) is the probability that the user has been observed (made a call) at Voronoi cell \( i \) subsequent to being observed at cell \( j \): if the individual was observed at \( j \) for a total of \( n_j \) times and subsequently at tower \( i \) with no observations in between for a total of \( n_{ij} \) times \((0 \leq n_{ij} \leq n_j)\) during the observation period, \( m_{ij} \equiv n_{ij}/n_j \). The matrix \( \mathbf{M} \) constructed in this manner exhibits the following mathematical properties. First, each column of \( \mathbf{M} \) sums to unity since \( \sum_i m_{ij} = \sum_i n_{ij}/n_j = 1 \). The same cannot be said for the rows, however, since \( \mathbf{M} \) is asymmetric in general. Second, from the Perron–Frobenius theorem, \( \mathbf{M} \) always has a leading eigenvalue \( \lambda_0 \equiv 1 \), whose associated right eigenvector is \( \mathbf{e}_0 = \mathbf{p} \), where \( \mathbf{p} = \{p_1, p_2, \ldots, p_K\} \) represents the stationary probability of observing the user at each tower location \((p_i = n_i/N)\). This can be easily understood: the stationarity condition is \( p_i = \sum_j m_{ij} p_j \), which we can also write as \( \mathbf{p} = \lambda \mathbf{M} \mathbf{p} \) with \( \lambda = 1 \). Lastly, all eigenvalues satisfy \(|\lambda_i| \leq 1\), i.e. they are confined within the unit circle on the complex plane.

Our presentation of human motion as a stochastic Markov process involves some simplification. Firstly, we assume that the calls were placed evenly spaced throughout the six-month observation period. Therefore, we do not address the impact of the inter-call time distribution [15] or the possible explicit time dependence of the element \( m_{ij} \) of \( \mathbf{M} \). Secondly, our \( \mathbf{M} \) incorporates the history dependence of one step only: a transition depends only on the current location, but not further into the past. One can certainly make the model more realistic by adding more detail to the model. We may, for instance, construct a time-dependent matrix \( \mathbf{M}(t) \) that reflects the changes in call patterns during the course of the day. We may also construct \( \mathbf{M} \) to be of a higher dimension (e.g., \( K \times K \times K \)) so that it incorporates history further into the past than one step, as we do here. But doing so would render the analysis presented in this paper unnecessarily complex. Nevertheless, our current construction of \( \mathbf{M} \) still yields interesting and useful insights to understanding human mobility, which we hope lays a foundation for future research that incorporates those possibilities.

3. Time-evolution of the radius of gyration

The range of human motion was quantified in González et al using the radius of the gyration \( r_g \), defined as [16]

\[
r_g^2(s) \equiv \frac{1}{s} \sum_{l=1}^{s} \left| \vec{r}(l) - \frac{1}{s} \sum_{k=1}^{s} \vec{r}(k) \right|^2 \quad (1 \leq s \leq N)
\]

a dynamic variable (as a function of \( s \), the index of the calls in figure 1(a)) that measures the overall range of an individual trajectory, which demonstrated the stark contrast between actual human motion and classical random walk models. We here ask whether

\[4\text{ In order to avoid the last cell visited at } t_{\text{max}} \text{ becoming a ‘sink’, which happens when there is no record of transition away from it, we assume that the individual makes an additional transition from it to the first Voronoi cell visited at } t_1.\]
The eigenmode analysis of human motion

Figure 2. Comparison of the evolution of the radii of gyration (normalized) between actual data (blue) and simulations using the Markov matrix $M$ (green) and the Bernoulli matrix $B$ (yellow). $M$ exhibits a markedly closer reproduction of the actual data, indicating that incorporating history (in the case of $M$, one step) is essential in modeling human mobility.

Our stochastic process approximation is able to reproduce its slow, sub-polynomial growth shown in [3]. In figure 2 we compare, with the x-axis rescaled from $s$ to real time $t$, the empirical temporal growth of $r_g(t)/r_g^{\text{final}}$ and that from simulations based on $M$, averaged over all individuals with $r_g^{\text{final}} \geq 100 \text{ km}$. We see that the actual curve and the simulation curve match within $\sim 5\%$, validating our approximation. To highlight the importance of incorporating history we also plotted the corresponding curve generated from the Bernoulli matrix $B$: a Bernoulli matrix is a transition matrix, but with no history dependence: its entry $b_{ij}$ is simply equal to $p_i = n_i/N$, the stationary probability to be at $i$ with no regard to one’s present location. $B$ shows a markedly quicker convergence of $r_g$ to $r_g^{\text{final}}$, significantly deviating from actual data. This handily demonstrates that temporal correlation between individual’s locations is an essential factor in modeling human mobility. In the following, we study the relationship between $r_g(s)$, equation (1) and the transition matrix to better understand the origin of this behavior.

3.1. Radius of gyration and the transition matrix

Under the Markov process $M$, the expected value of the square of the radius of gyration, equation (1)$^5$, at the $s$th step can be written as

$$\overline{r_g^2(s)} = \frac{1}{s} \sum_{l=1}^{s} \overline{r(l)^2} - \frac{1}{s^2} \sum_{l,k} \overline{r(l) \cdot r(k)} = \left(1 - \frac{1}{s}\right) (x^T P x + y^T P y)$$

$$- \frac{2}{s^2} \sum_{s'=1}^{s} (x^T M^{s'} P x + y^T M^{s'} P y), \quad (2)$$

where $\overline{r(l)}$ denotes the location of the user at the $l$th step ($\leq s$), $P = \{P_{ij} = p_{ij}\}$ is a diagonal $K \times K$ matrix, and $x = \{x_1, \ldots, x_K\}$ and $y = \{y_1, \ldots, y_K\}$ are the coordinates

$^5$ We study the squared form $r_g^2$ rather than $r_g$ which is much more complicated to treat analytically.

doi:10.1088/1742-5468/2010/11/P11021
of the $K$ unique Voronoi cells that a user has visited. Since in the stationary state $p_i$ represents the probability of finding the individual at cell $i$ at any step, the vector product $\vec{r}(l) \cdot \vec{r}(l + s')$ depends only the step difference $s'$, leading to $\vec{r}(l) \cdot \vec{r}(l + s') = \vec{r}(1) \cdot \vec{r}(1 + s') = x^T M^s' P x + y^T M^s' P y$.

The transition matrix can be expressed as $M = L^{-T} DR$, where $L$ and $R$ are its left- and right-eigenvector matrices normalized so that $RL^T = I$, and $D = \{ \delta_{ij} = \lambda_i \delta_{ij} \}$ is the diagonal matrix of $M$'s eigenvalues. Therefore we can rewrite equation (2) as

$$\overline{r^2_R}(s) = \sum_{k=1}^{K} (a_{x}^k b_{x}^k + a_{y}^k b_{y}^k) \left[ \left( 1 - \frac{1}{s} \right) - \frac{2}{s^2} \sum_{s'=1}^{s} (s - s') \lambda_k^{s'} \right]$$

$$= \sum_{k=1}^{K} \rho_k \left[ 1 - \left( \frac{1 + \lambda_k}{1 - \lambda_k} \right) \frac{1}{s} + \left( \frac{2 \lambda_k (1 - \lambda_k^2)}{(1 - \lambda_k)^2} \right) \frac{1}{s^2} \right]$$

$$\equiv \sum_{k} \rho_k \mathcal{F}(\lambda_k, s),$$

(3)

where we have defined the mode weight $\rho_k \equiv (a_{x}^k b_{x}^k + a_{y}^k b_{y}^k)$, invariant in $s$, with

$$R^{-1} P x = \{ a_{x}^k \} \quad \text{and} \quad x^T L^{-T} = \{ b_{x}^k \}^T$$

$$R^{-1} P y = \{ a_{y}^k \} \quad \text{and} \quad y^T L^{-T} = \{ b_{y}^k \}^T.$$  

(4)

Thus we have decomposed the expected value of $\overline{r^2_R}(s)$ into a sum of independent contributions from the $K$ eigenmodes of $M$. Specifically, equation (3) says that each eigenmode contributes in two ways, i.e. via the $s$-independent weight $\rho_k$ and the $s$-dependent $\mathcal{F}(\lambda_k, s)$. We show in figure 3 the behavior of $\mathcal{F}(\lambda, s)$: (a) when $\lambda$ is real (i.e. $\lambda \equiv \text{arg}(\lambda) = 0$ or $\pi$), the closer $\lambda$ is to $-1$, the faster the convergence of $\mathcal{F}(\lambda, s)$ to its maximum value 1. (Interestingly, $\lambda_0 = 1$ does not contribute at all, since $\mathcal{F}(1, s) \equiv 0$. Also, the Bernoulli matrix $B$ has $\lambda = 0$ as its only eigenvalue except the leading $\lambda_0 = 1$. Compare the yellow curve in figure 2 and the curve for $\lambda = 0$.) (b) When $\lambda$ is complex (i.e. $\lambda = \text{arg}(\lambda) \neq 0, \pi$), the convergence to the maximum is also accelerated as $\omega$ approaches $\pi$ for a given value of $|\lambda|$ (equal to 0.99 in figure 3(b)).

3.2. Eigenmode weight $\rho$  

Comparing figures 2 and 3, we find that the typical temporal behavior of $r_k$ is most likely due to modes with eigenvalues that are real and positive, close to one. Thus there are two possible scenarios: the first is that most eigenvalues of $M$ happen to be $\approx 1$, resulting in a slow increase of $r_k$ regardless of any other factors in equation (3) such as $x$ and $y$. The other is that, while the eigenvalues are broadly distributed over the complex unit circle, only those that are close to 1 carry large mode weights $\rho$.

To find the more likely scenario, for each individual we investigated two sets of eigenvalues, $S_1$ and $S_2$, defined as follows:

6 In figure 3 we show only the real portion of $\mathcal{F}$; in truth, $\text{Im}[\mathcal{F}] \neq 0$ when $\text{Im}[\lambda] \neq 0$, i.e. complex. In that case, its conjugate $\lambda^*$ is also an eigenvalue of $M$ with mode weight $\rho^*$. The combined effect of the two are basically $\rho \mathcal{F}(\lambda, s) + \rho^* \mathcal{F}(\lambda^*, s) = 2\text{Re}(\rho \mathcal{F}(\lambda, s))$.

doi:10.1088/1742-5468/2010/11/P11021
The eigenmode analysis of human motion

Figure 3. The increasing behavior of the real part of $F(\lambda, s)$ for various $\lambda$ values. (a) When $\lambda$ is real, $F(\lambda, s)$ increases slower as $\lambda \to 1$, while for $\lambda_0 \equiv 1$, $F(1, s) \equiv 0$. (b) The tendency is similar in the case of complex $\lambda$ (shown here for $\lambda = 0.99e^{i\omega}$), showing a faster convergence to one as $\omega$ is tuned away from zero.

- Set $S_1$ comprises eigenvalues with $|\lambda| \geq 0.6$ and $-10^\circ \leq \arg(\lambda) \leq 10^\circ$, contained within the blue and the yellow areas inside the unit circle in figure 4(a).
- Set $S_2$ comprises eigenvalues with $|\lambda| \geq 0.8$ and $-5^\circ \leq \arg(\lambda) \leq 5^\circ$, contained within the yellow area in figure 4(a).

If the second scenario is correct, we should see only a small fraction of eigenmodes contained within $S_1$ and $S_2$, while their combined mode weights $\sum_{k \in S} \rho_k$ account for most of $\sum_{k} \rho_k = r_g^2$. In fact, this is precisely what we see: in figure 4(b), we see that for 85% and 99% of the individuals, respectively, $S_1$ and $S_2$ contain fewer than 20% of their eigenvalues. On the other hand, figure 4(c) shows that, for 81% and 74% of the individuals, respectively, eigenmodes in $S_1$ and $S_2$ account for 80%–120% of $r_g^2$ (note that $\rho$ need not be positive for every eigenmode—it is the total sum of $\rho$ that is equal to $r_g^2$—so the sum of $\rho$ of a subset of eigenmodes may exceed $r_g^2$). Thus we can conclude that only a minority of eigenvalues near 1 account for most of $r_g^2$, causing the slowly increasing behavior we see in figure 2.

Next, we further explore the implications of such behavior at the individual level by investigating in detail the meaning of the eigenmodes of $M$.
The eigenmode analysis of human motion

Figure 4. (a) Eigenvalues of $M$ are distributed over the unit circle on a complex plane. We define two sets, $S_1$ and $S_2$, of eigenvalues close to 1 for each individual contained in the two colored areas inside the unit circle, excluding $\lambda_0 \equiv 1$. (Note that $S_1 \supset S_2$.) (b) Histogram of the fraction of an individual’s eigenvalues contained in $S_1$ and $S_2$. For 85% and 99% of individuals, $S_1$ and $S_2$ respectively contain fewer than 20% of their eigenvalues, indicating the broad distribution of eigenvalues over the complex unit circle. (c) Histogram of the fraction of total mode weights from eigenmodes in $S_1$ and $S_2$, $\sum_{k \in S} \rho_k / r_g^2$. For 81% and 73% of individuals, eigenmodes in $S_1$ and $S_2$ account for 80%–120% of $r_g^2$, indicating their disproportionate significance in the dynamics of the radius of gyration.
4. The meaning of eigenmodes

Let $p_k(s)$ denote the probability to find an individual at a Voronoi cell $k$ at step $s$ under the Markov process. Then we can express $p(s) = (p_1(s), p_2(s), \ldots, p_K(s))$ as

$$p(s) = M^s p(0) = \sum_{k=1}^{K} a_k \lambda_k^s e_k,$$

where $\{e\}$ are the right eigenvectors of $M$. The coefficients $\{a_k\}$ can be obtained from $R^{-1} p(0) = (a_1, a_2, \ldots, a_n)$. This means that $p(s)$ is a sum of independent temporal evolutions of the eigenmodes themselves. Note that $a_k$, $\lambda_k$, and $e_k$ can be complex due to the asymmetric nature of $M$. A complex-valued eigenmode would contribute to $p(s)$, in conjunction with its complex conjugate (which is also an eigenmode of $M$), as

$$a_k \lambda_k^s e_k + a_k^*(\lambda_k^*)^s e_k^* \propto |a_k| |\lambda_k|^s (\ldots, |e_{k,n}| \cos(\alpha_k + \omega_k s + \phi_{k,n}), \ldots),$$

where $e_{k,n}$ is the $n$th component of $e_k$, and $\alpha_k \equiv \arg(a_k)$, $\omega_k \equiv \arg(\lambda_k)$, and $\phi_{k,n} \equiv \arg(e_{k,n})$, guaranteeing that $p(s)$, a probability, is always real.

In the asymptotic $s \to \infty$ limit all modes decay to zero except the leading mode with $\lambda_0 \equiv 1$, so that $p(\infty) = p = \{p_i\}$, the stationary occupation probability. For finite $s$, however, the transient dynamics described by non-leading eigenvalues is important and offer interesting detail of the physical process. In this section, therefore, we study the non-leading eigenmodes, and discuss how they can characterize individual human mobility patterns. For convenience, we discuss real- and complex-valued eigenmodes separately.

4.1. Real-valued eigenmodes

Note that equation (6) holds for a real-valued eigenmode as well: the angular variables $\{\alpha_k, \omega_k, \phi_{k,n}\}$ are merely restricted to either 0 or $\pi$. First, consider $\phi_{k,n}$. Within the same eigenmode $k$, it is possible that some $\phi_{k,n}$ are 0 while others are $\pi$ (equivalent to the components being positive or negative). Since each component corresponds to a Voronoi cell, this means that we can cluster the Voronoi cells into two distinct groups, according to the sign of their eigenvectors (only in regards to this specific eigenmode; other eigenmodes can offer different groupings). According to equation (6), this means that as $s$ is incremented, the occupation probabilities of Voronoi cells in the same cluster are in phase, and out of phase with cells in the other cluster. Additionally, if $\lambda_k > 0$ ($\omega_k = 0$), the occupation probabilities attenuate monotonically, while if $\lambda_k < 0$ ($\omega_k = \pi$), they fluctuate as they attenuate. This behavior is visualized in figure 5(a), showing the occupation probabilities of the two said clusters (blue and yellow).

We can rephrase this in a more intuitive way: When $\lambda_k > 0$ ($\omega_k = 0$), the individual is likely to be contained within the same cluster as time progresses, with a small probability of transitioning into the other cluster. If the individual does cross over to the other cluster, however, the trend persists: their motion is now likely to persist in that cluster of Voronoi cells. On the other hand, when $\lambda_k < 0$ ($\omega_k = \pi$), at every step the individual tends to transition between clusters, resulting in a large intercluster probability of flux exchange, hence the fluctuation.

\[\alpha_k,\] dependent only on the initial condition, merely dictates the relative magnitude of each eigenmode in equation (6) at $s = 0$. 

\[\text{doi:10.1088/1742-5468/2010/11/P11021} \]
Figure 5. (a) The two cases of the temporal evolution of the occupation probabilities in eigenmodes with real $\lambda$. When $\lambda > 0$ ($\omega = 0$), the occupation probability of a Voronoi cell attenuates monotonically towards the stationary state, meaning a small transition between the different clusters (blue and yellow). When $\lambda < 0$ ($\omega = \pi$), the occupation probabilities fluctuate sinusoidally, indicating an active exchange of probability flux between the two clusters (the gray envelope curve is a guide). (b) A simple Markov process consisting of three cells with symmetric transition probabilities, along with a diagrammatic summary of the dynamic properties of its eigenmodes given in equation (8). In eigenmode 1, cell 1 constitutes one cluster (blue) and cells 2 and 3 the other cluster (yellow). In accordance with (a), when $\lambda_1 > 0$, the individual’s motion is mostly confined to inside one cluster, whereas when $\lambda_1 < 0$, it transitions often between the two clusters. In eigenmode 2, cells 2 and 3 constitute two distinct clusters (cell 1 is irrelevant, with its eigenvector component 0 in $e_2$). Thus eigenmode 2 presents a finer detail of motion—between cells 2 and 3—which is absent in eigenmode 1. (c) A real-world example of Voronoi cell clustering. A portion of a major highway is contained within ten Voronoi cells (left). According to the eigenmode 1 ($\lambda_1 > 0$) of an individual who traveled along these cells, the entire portion of the highway belongs to a single cluster. However, this does not tell us whether they were traversed in a linear fashion or in a haphazard way (however unlikely). According to the eigenmode 2, however, the highway is divided into two adjacent compact clusters. In accordance with (b), $\lambda_2 > 0$ tells us that there were minimal transitions between the clusters, indicating that the cells were traversed in a linear fashion, which agrees with our intuition.
The eigenmode analysis of human motion

We can illustrate this using a simple construction given in figure 5(b). In it we have three Voronoi cells represented as nodes 1, 2, and 3. The corresponding transition matrix $M$ is

$$
M = \begin{pmatrix}
1 - 2P & P & P \\
P & 1 - P - Q & Q \\
P & Q & 1 - P - Q
\end{pmatrix},
$$

which results in the following three eigenmodes:

Eigenmode 0 : $\lambda_0 = 1$, 
$$
e_0 = \frac{1}{3}(1, 1, 1)
$$

Eigenmode 1 : $\lambda_1 = 1 - 3P$, 
$$
e_1 = \frac{1}{6}(2, -1, -1)
$$

Eigenmode 2 : $\lambda_2 = 1 - P - 2Q$, 
$$
e_2 = \frac{1}{2}(0, 1, -1).
$$

Asymptotically ($s \to \infty$), an individual has an equal probability $p = e_0 = (1/3, 1/3, 1/3)$ to be found at each cell. Eigenmodes 1 and 2 dictate the transient behavior: According to equations (8), in eigenmode 1, cell 1 has a positive eigenvector component ($\phi_{1,1} > 0$) while 2 and 3 have negative components. Thus if $\lambda_1 = 1 - 3P > 0$, the walker tends to be confined at cell 1 or in the cluster of cells 2 and 3. This is intuitively understandable; a small $P$ means that transition between the two clusters is discouraged, leading to the confinement, and vice versa. Eigenmode 1 does not give us the details on the transition between 2 and 3, however. This finer detail is given by eigenmode 2: In it, cells 2 and 3 are in separate clusters (cell 1 is completely irrelevant with $e_{2,1} = 0$): if $\lambda_3 = 1 - P - 2Q > 0$, the walker is confined to either 2 or 3 with only a small probability of transitioning to the other, and vice versa. This is presented graphically in figure 5(b).

The discussion so far puts us in a position to explain the behaviors of $F(\lambda, s)$ we see in figure 3. When $\lambda < 0$, the individual has a high propensity to transition to the other cluster than one he is in. Thus the individual covers a larger area in the same time, meaning that the radius of gyration increases faster than in the case of $\lambda > 0$. Therefore, concentration of mode weights $\rho$ on modes with $\lambda \approx 1$ is the manifestation of the nature of relatively local movements of humans—were it the opposite, we should see large $\rho$ on modes that show a faster increase of the radius of gyration.

We present a real-world example of the Voronoi cell clustering taken from a specific individual’s data in figure 5(c). It shows a small rectangular portion of the country containing ten Voronoi cells (left). A segment of the nation’s highway runs along the set. We consider two largest eigenmodes, 1 and 2, of an individual who has traversed the cells. $\lambda_1$ and $\lambda_2$ satisfy $1 > \lambda_1 > \lambda_2$. Similar to the simple example of figure 5(b), according to eigenmode 1 the entire strand forms a single cluster (center). This agrees well with our intuition regarding a traveler on a highway—once they enter a highway, it is likely that they stay on it. Yet, this is not yet a conclusive proof that the cells were visited in succession; we cannot exclude the possibility, however bizarre it may sound, that the user had driven in a haphazard manner along the cells. The cell clustering according to eigenmode 2 (right) provides an answer: $\lambda_2 > 0$ means that there could have been only minimal transitions between the two clusters, showing that the traveler did move in a linear fashion on the highway.

8 The cells have been slightly deformed to mask the identity of the country.
4.2. Complex-valued eigenmodes

We can readily extend the interpretation of real eigenmodes of section 4.1 into that of complex-valued eigenmodes; we need to consider that now the angular variables \( \{ \alpha_k, \omega_k, \phi_{k,n} \} \) in equation (6) can be of any value in \([0, 2\pi)\). This means that the Voronoi cells can be grouped into more than two clusters, and that the occupation probabilities of each cell may oscillate with varying periods. We again illustrate this point with a simple example shown in figure 6(a), which exhibits a natural 120° rotational symmetry. Solving for its transition matrix, we obtain the following eigenmodes (without loss of generality, we assume \( P > Q \)):

\[
\text{Eigenmode 0 : } \lambda_0 = 1, \quad e_0 = \frac{1}{3}(1, 1, 1)
\]

\[
\text{Eigenmode 1 : } \lambda_1 = \frac{1}{2}((2 - 3(P + Q)) + i\sqrt{3}(P - Q)), \quad e_1 = (e^{-i2\pi/3}, e^{i2\pi/3}, 1)
\]

\[
\text{Eigenmode 2 : } \lambda_2 = \lambda_1^*, \quad e_2 = e_1^*, \quad (9)
\]

The asymptotic occupation probability is again \( p = e_0 = (1/3, 1/3, 1/3) \). More interestingly, we observe that the eigenvector \( e_1 \) and its conjugate \( e_2 \) cluster the cells into three distinct groups, regardless of \( P \) and \( Q \). The transient dynamics of the traveler determined by \( \omega_1 \) and \( \omega_2 \), however, does not always reflect the underlying 120° symmetry (i.e., a rotational movement of period three). To see why, let us rewrite \( |\lambda_1| \) and \( \omega_1 = \arg(\lambda_1) \):

\[
|\lambda_1| = \frac{1}{2}\sqrt{(2 - 3\delta)^2 + 3\eta^2}, \quad \omega_1 = \arctan \left( \frac{\sqrt{3}\eta}{2 - 3\delta} \right), \quad (10)
\]

where we have defined \( \delta \equiv (P + Q) \) and \( \eta \equiv (P - Q) \). We can say that \( \delta \) represents the dynamism of the traveler since it is the probability to transition to a different cell at each step, while \( \eta \) represents the rotational tendency of the traveler since a large \( \eta \) would result in a stronger clockwise tendency of movement. First, assume \( P = 1 \) and \( Q = 0 \) so that \( \delta = \eta = 1 \). This leads to \( |\lambda_1| = 1 \) and \( \omega_1 = 2\pi/3 \), and resulting from it is a perpetual (non-attenuating) clockwise rotational motion of period three, perfectly reflecting the symmetry of the diagram. Any other combination of \( P \) and \( Q \) leads to a slower, attenuating motion with larger periods, if at all—from the construction of figure 6(a), a non-zero \( Q \) or \( 1 - P - Q \) would lead to what would function as ‘friction’ against a pure rotation of period three.

Yet, given that a period-three rotation is the simplest periodic motion arising from complex-valued eigenmodes, it appears worth investigating whether such motion can be readily found in real human trajectories. To check the possibility, we studied the distribution of \( \omega = \arg(\lambda) \) for eigenvalues with magnitude \( |\lambda| \geq 0.5 \) (for reasonable persistence), shown in figure 6(b). We find that, although the majority of eigenvalues are real \( (\omega = 0 \text{ or } \pi) \), there exist prominent peaks at \( \omega \approx 2\pi/3 \approx 120° \). In fact, the peaks represent a sizable portion of the population (in our dataset, 24.7% of the individuals)\(^9\). Although our current data set does not include detail on the exact places of visit or activities undertaken by each individual, we suggest a potentially testable movement pattern that appears to be a reasonable origin of the behavior in

\(^9\) Although it would not be an isolated movement like the one shown in figure 6(a), it means that it is sufficiently isolated to exhibit such eigenvalues.
Figure 6. (a) A simple asymmetric Markov process with a natural 120° rotational
symmetry. According to eigenmodes 1 and 2 (complex conjugates of each other) in equations (9), each cell constitutes a distinct cluster, resulting in three separate clusters (blue, yellow, and green in the figure). The dynamic property of the motion depends on $P$ and $Q$. When $(P, Q) = (1, 0)$, a permanent (non-attenuating) clockwise rotation of period three occurs. In more general cases of $(P, Q) \neq (1, 0)$, a combination of stationary $(1 - P - Q)$ and dynamic $(P + Q)$ tendencies lead to an attenuating rotation of varying prolonged periods. (b) The distribution of $\omega$ for eigenmodes with $|\lambda| > 0.5$. In addition to real eigenvalues ($\omega \in \{0, 180°\}$), we find prominent peaks at $\omega \approx 2\pi/3(120°)$, indicating that a rotational motion of (nearly) period three is commonplace. (c) An example trajectory giving $\omega \approx 2\pi/3$ as a prominent eigenmode of motion.
5. Discussion

In this paper, we have presented in detail how the framework of stochastic processes, widely used in theoretical physics, can be used for analysis of large-scale human motion data. Specifically, utilizing the well-established theory of Markov matrices, we have demonstrated that the observed temporal evolution of the radius of gyration can be understood via eigenmode analysis of individual transition matrices. We have also discussed how the eigenvalues and eigenvectors are related to the microscopic characteristic modes of individual mobility.

We anticipate our approach to grow more relevant as innovations continue in large-scale data acquisition technology: even now, with GPS-enabled mobile phones becoming more available, individuals are able to track their own movements over locales of interest and couple them with digitalized geographical information (the so-called ‘geo-tagging’) to construct a detailed space–time history of one’s past whereabouts. On a more social scale, furthermore, we can readily imagine the benefits of a better understanding of human mobility patterns: it would allow us to better design infrastructure such as roads, transportation systems, and vital utilities so that social cost is minimized while location-based human activities are optimally supported [17]. Also notable is the active research effort in the field of ecology to understand animal movements [18,19] which have a potential to be helpful in understanding human mobility in urban environments as well.

We hope that our work plays a role in highlighting the opportunities for theoretical physicists to make novel and innovative contributions to social and technological problems.

Acknowledgments

We would like to thank A-L Barabási and Doochul Kim for helpful suggestions. This work was supported by Kyung Hee University grant KHU-20100116, the National Research Foundation of Korea grant KRF-20100004910, the NRF grant funded by MEST (2010-0015197) and the T J Park Junior Faculty Fellowship.

References

[1] Editorial, Community cleverness required, 2008 Nature 455 1
[2] Brockmann D, Hufnagel L and Geisel T, 2006 Nature 439 462
[3] González M, Hidalgo C and Barabasi A-L, 2008 Nature 453 779
[4] Mantegna R N and Stanley H E, 1994 Phys. Rev. Lett. 73 2946
[5] Kläfter J, Shlesinger M F and Zuevomen G, 1996 Phys. Today 49 33
[6] Hughes B D, 1995 Random Walks in Random Environments vol 1 (Oxford: Oxford University Press)
[7] Hughes B D, 1996 Random Walks in Random Environments vol 2 (Oxford: Oxford University Press)
[8] Sinai Y G, 1983 Theory. Probab. Appl. 27 256
[9] Durrett R, 1986 Commun. Math. Phys. 104 87
The eigenmode analysis of human motion

[10] Kesten H, 1986 Physica A. 138 299
[11] Markov A A, 2006 Sci. Context 19 591
[12] Shannon C E, 1948 Bell Syst. Tech. J. 27 379

Shannon C E, 2001 ACM SIGMOBILE Mobile Comput. Commun. Rev. 5 3
[13] Bremaud P, 1991 Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues (New York: Springer)
[14] Krogh A, Brown M, Mian I, Sjolander K and Haussler D, 1994 J. Mol. Biol. 235 1501
[15] Candia J, González M, Wang P, Schoenhart T, Madey G and Barabasi A-L, 2008 J. Phys. A: Math. Theor. 41 224015
[16] Goldstein H, 1980 Classical Mechanics (Reading: Addison-Wesley)
[17] Um J, Son S, Lee S, Jeong H and Kim B, 2009 Proc. Natl. Acad. Sci. 106 14236
[18] Nathan R, Getz W, Revilla E, Holyoak M, Kadmon R, Saltz D and Smouse P, 2008 Proc. Natl. Acad. Sci. 105 19052
[19] Revilla E and Wiegand T, 2008 Proc. Natl. Acad. Sci. 105 19120

doi:10.1088/1742-5468/2010/11/P11021