

Measuring regularity of human physical activities

Keqin Shi, Zhen Chen, Weiqiang Sun, *Senior Member, IEEE*, and Weisheng Hu, *Member, IEEE*

Abstract—Regularity is an important aspect of physical activity that can provide valuable insights into how individuals engage in physical activity over time. Accurate measurement of regularity not only advances our understanding of physical activity behavior but also facilitates the development of human activity modeling and forecasting. Furthermore, it can inform the design and implementation of tailored interventions to improve population health outcomes. In this paper, we aim to assess the regularity of physical activities through longitudinal sensor data, which reflects individuals' all physical activities over an extended period. We present several entropy models, including entropy rate, approximate entropy, and sample entropy, which offer a more comprehensive evaluation of physical activity regularity compared to metrics based solely on periodicity or stability. We also propose a framework to validate the performance of entropy models on both artificial and real-world physical activity data. The results indicate entropy rate is able to identify not only the magnitude and amount of noise but also macroscopic variations of physical activities, such as differences on duration and occurrence time. Simultaneously, entropy rate is highly correlated with the predictability of real-world samples, further highlighting its applicability in measuring human physical activity regularity. Leveraging entropy rate, we investigate the regularity for numerous individuals. We find the composition of physical activities can partially explain the difference in regularity among individuals, and the majority of individuals exhibit temporal stability of regularity.

Index Terms—Physical activity, regularity, entropy rate, approximate entropy, sample entropy

I. INTRODUCTION

PHYSICAL activity is defined as any voluntary bodily movement produced by skeletal muscles that requires energy expenditure [1]. Physical activity encompasses all activities, at any intensity, performed during any time of day or night. It includes both exercise and incidental activity integrated into daily routines, such as walking to the local shop, cleaning, working, active transport, etc [2]. Driven by daily routine, some physical activities of humans over a period of time are regular, for instance, commuting on time during weekdays or exercising regularly. On the other hand, influenced by the external environment and unexpected events, the physical activities of humans often have the random nature as well. To what extent are an individual's physical activities

regular? This is a fundamental question about human behavior understanding. Our goal here is to measure the regularity of human physical activities. The accurate measurement of regularity can facilitate advancements in human activity modeling and prediction, and further enable the implementation of tailored interventions aimed at improving health outcomes.

Conventional methods [3], [4] measure the regularity of human life by designing diary-like instruments to record the occurrence of particular events. Such data collection requires the subjects to record their activities of daily living manually. Therefore it is expensive and subjective, and also constrained to small sample sizes and short time spans. However, the proliferation of smartphones and wearable devices provides new opportunities to collect sensor data about human physical activities on a large scale and over a long period of time. The sensor data collected from smartphones and wearable devices includes step counts, calorie expenditures, exercise intensities, heartbeats, and so on. Such longitudinal sensor data sampled at intervals of a few minutes can reflect an individual's all physical activities more accurately and informatively [5]. This makes it possible to measure the regularity of human physical activities in detail and reliably.

In the existing literature, some metrics have been proposed to measure the regularity of human physical activity utilizing longitudinal sensor data. Generally, these metrics can be divided into two categories: periodicity-based metrics and stability-based metrics. The former focus on measuring the 24-hour periodicity of sensor data. These periodicity-based metrics calculate the similarity of data points at 24-hour intervals as an indicator of regularity, and they disregard the orderliness of successive data points. For the latter case, regularity is denoted by stability. These metrics use the standard deviation of the original sensor data or some features extracted from the sensor data to determine regularity. Despite this, sensor data with large variations can also be considered regular, such as periodic signals. Both periodicity and stability are special types of regularity. As these metrics presuppose types of regularity in advance, they cannot provide a comprehensive measurement of regularity for human physical activity.

Regularity refers to the extent to which individual activities repeat over time in fixed patterns [6]. It not only depends on the variability in the characteristics of a single type of activity but also on the cycles or patterns existing in activity sequences. In information theory, entropy quantifies the uncertainty of a random variable and it also measures the degree of randomness in the system [7]. There are several entropy models that can be used to determine the regularity of serial data based on the presence of patterns. The entropy rate measures how the entropy of a sequence changes over

This paragraph of the first footnote will contain the date on which you submitted your paper for review. This research is supported by the National Key Research and Development Project of China (2019YFC1709800).

Keqin Shi, Zhen Chen, Weiqiang Sun and Weisheng Hu are with the Department of Electronic Engineering, Shanghai Jiao Tong University, 800 Dongchuan Rd. 200240, Shanghai (e-mail: keqinshi@sjtu.edu.cn; 838301834@sjtu.edu.cn; sunwq@sjtu.edu.cn; wshu@sjtu.edu.cn).

time, allowing the randomness of the sequence to be quantified [8]. In practice, the entropy rate has been used extensively to assess the regularity of categorical sequences, including human mobility [6], [9], human economic behavior [10], human online life [11], web browsing behavior [12] and patient health records [13]. Additionally, Pincus *et al.* [14] devised Approximate Entropy (ApEn) as the measurement of regularity that can be applied to short and noisy time series of continuous values. Richman *et al.* [15] developed Sample Entropy (SampEn), which addresses the issues of bias and lack of relative consistency in approximate entropy, to measure the regularity of clinical and experimental time series. While approximate entropy and sample entropy are initially developed for physiological applications, both of them have been used in other fields such as medicine [16]–[18], economics [19], [20], climatology [21]–[23], gait analysis [24], [25], and battery health prognostics [26], [27]. Despite this, there is, to date, a paucity of research on the application of these entropy models to human physical activity data. It still remains unclear which entropy model is most suitable for measuring the regularity of human physical activities.

In this paper, we aim to quantify the regularity of human physical activities accurately and comprehensively from longitudinal sensor data. First, the entropy rate calculation procedure is modified to allow it to be applied to physical activity data with continuous values. Then, we compare the applicability of entropy rate, approximate entropy, and sample entropy as measurements of the regularity of physical activities. By simulating real-life activity patterns with artificial data, we validate the performance of entropy models under different scenarios. Results indicate that entropy rate is superior to approximate entropy and sample entropy. Entropy rate can identify not only the magnitude and amount of randomness, but also the macroscopic variations, such as the differences on duration and occurrence time. We then further evaluate the performance of the three entropy models by correlating their respective entropy values with prediction errors obtained from multiple forecast models, on real-world physical activity data samples. We show that again entropy rate outperforms the other two entropy models with a correlation coefficient as high as 0.895. We then conclude that entropy rate is a reliable measurement of human physical activity regularity.

Utilizing entropy rate as the measurement of human physical activity regularity, we investigate the interpersonal and intrapersonal variation of regularity for 686 individuals. We find that the regularity varies considerably across individuals and the difference in activities composition can explain a large part of the variation. Meanwhile, the majority of individuals maintain stable physical activity habits and their regularity does not change significantly over time. The contributions of this study are summarized as follows:

- 1) We modify the calculation procedure of entropy rate so that it can be applied to physical activity data.
- 2) We propose a framework to validate the performance of entropy models on both artificial and real-world physical activity data. Experiment results demonstrate that entropy rate is more suitable than approximate entropy and sample entropy for measuring the regularity

of human physical activities.

- 3) Our analysis of human physical activity regularity using entropy rate reveals that variations in regularity among individuals are primarily associated with the composition of activities. In addition, the regularity of most individuals remains stable over time.

The rest of the paper is organized as follows. We first present the related work on human behavior regularity in section II. Then we describe the details on entropy rate, approximate entropy, and sample entropy in section III. Section IV displays the validation of three entropy models. More results about human physical activity regularity are shown in section V. Finally, in section VI, we conclude our work and discuss future research directions.

II. RELATED WORK

The concept of regularity is recognized as a fundamental aspect in the field of human behavior understanding. Prior to the widespread adoption of wearable technology, a significant body of literature has utilized survey data to assess the regularity in human daily activities. The Social Rhythm Metric (SRM) [3], [4] is a widely used metric that quantifies the regularity of an individual's daily activities with respect to their timing. The SRM is calculated by first having subjects record the occurrence of various event categories, and then determining the habit time of each event through an outlier elimination algorithm. The average count of events that occur within the habit time for each event category is then used to reflect the subject's level of daily lifestyle regularity. Despite its widespread usage in studies exploring the relationship between health outcomes and lifestyle regularity [28]–[30], the SRM ignores the interconnections among successive events, leading to an incomplete representation of regularity.

In recent years, mobile sensing has shown increasing potential for tracking human daily activities of living. The widespread use of smartphones and wearable devices has enabled the collection of rich, longitudinal sensor data that characterizes an individual's activities, sleep patterns, and application usage. These data provides a foundation for studying the regularity of human daily activities and has resulted in the proposal of several metrics. These metrics can be broadly categorized into two groups: those that focus on quantifying the periodicity in the sensor data and those that employ stability to capture the essence of regularity.

Since human behaviors are driven by an internal biological clock that regulates the sleep–wake cycle and repeats roughly every 24 hours, the 24-hour periodicity is usually regarded as a measurement of human life regularity. Saeb *et al.* [31] and Wang *et al.* [32] proposed circadian rhythm to measure the strength with which an individual follows a 24-hour rhythm in behaviors. Sensor data was converted to the frequency domain, and the circadian rhythm was determined by energy that falls within the 24 ± 0.5 h. Phillips *et al.* [33] developed the sleep regularity index (SRI), which calculates the percentage probability of an individual being in the same state (asleep vs. awake) at any two time-points 24 hours apart, as the measurement of sleep regularity. A similar metric can be

extended to sensor data other than sleep. In reference [32], [34], regularity index (RI) and flexible regularity index (FRI) were proposed to assess the difference between the same time points across different days. The RI calculates the product of rescaled values of the same time points from different days to evaluate the difference, while the FRI uses edit distance. It should be noted that these metrics only focus on the similarity of data points at 24-hour intervals and do not take into account the orderliness of successive data points.

Other metrics in the literature treat the stability as a proxy for regularity. Standard deviations are commonly used to measure the variance of daily activities. Marschollek *et al.* [35] employed the standard deviation of all time differences between physical activity event starts as a measure of regularity. Fischer *et al.* [36] applied the standard deviation to features of daily sleep (e.g. sleep onset, sleep offset, midsleep, duration) to quantify sleep regularity. Wang *et al.* [32] directly calculate standard deviation of physical activity data as a metric for human physical activity variability. Wil *et al.* [37] introduced the concepts of inter-daily stability and intra-daily variability to assess rest-activity rhythms. Inter-daily stability evaluates the consistency of daily activity patterns with respect to the average pattern across days, and it reflects the stability of rest-activity rhythms over multiple days. Intra-daily variability is calculated as the ratio of the mean squared the first deviation of the data and the population variance of the data, which detects fragmentation of rest-activity rhythms. In addition to rest-activity rhythms, they can also be applied to measure sleep regularity [36]. However, stability is only one aspect of regularity, as periodic signals with large variations can still be considered regular. Therefore, the use of stability alone may not be sufficient to represent regularity.

It is evident that none of the metrics previously discussed are capable of assessing the regularity of human daily activities comprehensively. These metrics are limited in their scope, as they make prior assumptions about the definition of regularity. A comprehensive regularity measurement should consider all possible patterns present in the sensor data. In the following section, we introduce entropy-based metrics that circumvent the limitations of the previously mentioned metrics, offering a more inclusive measurement of regularity.

III. MEASUREMENT OF REGULARITY

The antithesis of regularity is randomness. The analysis of the randomness of a series has its roots in information theory and the concept of entropy. Entropy quantifies the amount of information of random variables based on the probability distributions. It can also measures the degree of randomness in the system [8]. However, Shannon entropy has a limitation in its inability to capture patterns present in sequential data, as it disregards the temporal correlation among elements in sequence. The real entropy of a sequence depends not only on the frequency of elements in the sequence, but also on the order in which the elements are combined. Entropy rate, approximate entropy, and sample entropy are three kinds of entropy that consider the ordered sub-sequence existed in the sequence. And they are widely used to measure the regularity of sequential data.

A. Entropy rate

Mathematically, a series from real-world can be modeled as a stochastic process \mathbf{X} , which is an indexed sequence of random variables $[X_1, \dots, X_n]$. And there can be an arbitrary dependence among the random variables. The joint entropy of the collection of random variables is

$$H(X_1, \dots, X_n) = - \sum_{x_1 \in \Omega_1} \dots \sum_{x_n \in \Omega_n} p(x_1, \dots, x_n) \log_2 p(x_1, \dots, x_n). \quad (1)$$

The entropy rate of a stochastic process is the asymptotic rate at which the entropy of a sequence grows with increasing n . The entropy rate $H(\mathbf{X})$ is defined as follows.

$$H(\mathbf{X}) = \lim_{n \rightarrow \infty} \frac{1}{n} H(X_1, \dots, X_n). \quad (2)$$

As shown in equation (2), entropy rate is the average entropy over all random variables and it means the average information gain with the increment of the sequence. Reference [7] proves that this limit of equation (2) exists for all stationary random processes and is equal to

$$H(\mathbf{X}) = \lim_{n \rightarrow \infty} H(X_n | X_1, \dots, X_{n-1}), \quad (3)$$

where $H(X_n | X_1, \dots, X_{n-1})$ is the conditional entropy of the last variable given the previous $n - 1$ values. Equation (3) indicates that the entropy rate account for the dependencies among random variables. The stronger dependencies among variables in a stochastic process, the more information the previous variables provide about the next one, and therefore the lower the entropy rate of the process. In contrast, if all variables of the process are independent, the entropy rate of the process is exactly equivalent to the Shannon entropy of the process, which is the upper bound for the entropy rate.

The estimation of entropy rate can be challenging since it is difficult to know the joint probability distribution of finite sequences in real-world data. Here, we introduce an estimation algorithm based on Lempel-Ziv data compression [38], which is known to rapidly converge to the real entropy rate of a time series. For a time series with length n , the entropy rate is estimated by

$$H^{est} = \frac{n \times \log_2(n)}{\sum_{i=1}^n \Lambda_i}, \quad (4)$$

where Λ_i is the length of the shortest substring starting at position i which doesn't previously appear from position 1 to $i - 1$. We illustrate the estimation procedure of entropy rate of a discrete sequence $\mathbf{X} = (A, B, A, B, C)$ in Table I for a better understanding. The notation $\mathbf{X}[1 : i]$ is historical subsequence before position i . S_i is the shortest subsequence that never appeared in $\mathbf{X}[1 : i]$, and Λ_i is the length of the shortest subsequence S_i . For $i = 1, 2, 5$, $\Lambda_i = 1$, since the symbols in these positions are new symbols. While, for $i = 3$, the shortest new subsequence is ABC , because the historical subsequence AB appears in position 3 again. A similar situation also occurs in position 4. Above example tells us if there are some fixed patterns that appear in the sequence repeatedly, their Λ_i will be larger and the entropy rate will be smaller. In the extreme case of a sequence whose symbols are all unique, $\Lambda_i = 1$ for all symbols and $\sum_{i=1}^n \Lambda_i = n$. In this case, the entropy rate is

TABLE I

AN EXAMPLE ILLUSTRATING THE ESTIMATION OF ENTROPY RATE.

i	$\mathbf{X}[1:i]$	S_i	Λ_i
1	ABABC	A	1
2	ABABC	B	1
3	ABABC	ABC	3
4	ABABC	BC	2
5	ABABC	C	1

maximum, equal to the Shannon entropy of sequence, which is $\log_2(n)$. It also have been proven that H^{est} will converge to the actual entropy rate when n approaches infinity [39].

B. Approximate entropy

Approximate entropy is a statistic quantifying the regularity and complexity of short and noisy time series data [40]. It originated from the analysis of complexity in dynamic systems and is seen as the information-theoretic rate of entropy for approximating Markov chains [41]. Approximate entropy measures the logarithmic probability that nearby pattern runs remain close in the next incremental comparison. The calculation of approximate entropy requires two parameters, which are m , the length of the template, and r , a noise filter. Statistically, it would be the equivalent of dividing the space of states into cells of width r , to estimate the conditional probabilities of the m -th order.

Given a sequence of data $\mathbf{X} = [x_1, x_2, \dots, x_n]$ of length n , form subsequences $s_m(1)$ through $s_m(n-m+1)$, defined by $s_m(i) = [x_i, x_{i+1}, \dots, x_{i+m-1}]$. The distance between $s_m(i)$ and $s_m(j)$ is defined as the maximum difference in their respective scalar components, which is

$$d[s_m(i), s_m(j)] = \max_{k=1,2,\dots,m} (|x_{i+k-1} - x_{j+k-1}|). \quad (5)$$

We define a quantity named correlation integral,

$$C_i^m(r) = \frac{1}{n-m+1} \sum_{j=1}^{n-m+1} \#d[s_m(i), s_m(j)] \leq r, \quad (6)$$

which is the average number of subsequences similar to $s_m(i)$. Then we compute,

$$\Phi^m(r) = \frac{1}{n-m+1} \sum_{i=1}^{n-m+1} \log_2(C_i^m(r)). \quad (7)$$

Finally, the approximate entropy of the sequence is

$$\begin{aligned} ApEn(m, r) &= \Phi^m(r) - \Phi^{m+1}(r) \\ &\approx -(n-m)^{-1} \sum_{i=1}^{n-m} \log_2\left(\frac{C_i^{m+1}(r)}{C_i^m(r)}\right). \end{aligned} \quad (8)$$

From equation (8), we can see that approximate entropy is inversely proportional to the conditional probability that similar subsequences of length m stay consistent at the next position. Greater likelihood of remaining close, implying regularity, produces smaller $ApEn$ values, and conversely.

Pincus showed that approximate entropy would converge to the entropy rate for independent identical distribution series and finite Markov chains [14]. However, this does not hold in

more general cases, since approximate entropy is designed as a relative measurement used to compare the regularity of different time series. Approximate entropy can vary significantly with the choice of m and r , but the relativity of approximate entropy is enough to discriminate different systems. Pincus [42] pointed that, in general, given two data series X_1 and X_2 , when $ApEn(m_1, r_1)(X_1) < ApEn(m_1, r_1)(X_2)$ then $ApEn(m_2, r_2)(X_1) < ApEn(m_2, r_2)(X_2)$. Additionally, the selection of appropriate values for m and r is crucial in ensuring accurate estimation of the conditional probability from data series of length n . It is recommended that m has a relatively low value, e.g., 2 or 3, since a reasonable estimation of conditional probability needs preferably 30^m points. The value of r should fall within the range of $0.1\sigma - 0.25\sigma$, where σ is the standard deviation of series.

C. Sample entropy

Approximate entropy is biased statistic. The bias arises from the calculation of correlation integral $C_i^m(r)$, which allows each subsequence to count itself to ensure the logarithms remain finite. As a consequence, the conditional probability is overestimated. If we call B_i is the number of subsequences with length m that are similar to subsequence $s_m(i)$, and A_i is the number of subsequences with length $m+1$ that are similar to subsequence $s_{m+1}(i)$. The approximate entropy calculate $(A_i + 1)/(B_i + 1)$ as conditional probability, which is greater than the real one A_i/B_i . This bias is obviously more important for series with a small number of points n .

Richman *et al.* [15] defined sample entropy, a statistic which does not have self-counting and eliminates the bias of approximate entropy. And the calculation procedure of sample entropy is simpler than approximate entropy. We define:

$$A = \sum_{i=1}^{n-m} \sum_{j=1, j \neq i}^{n-m} \#d[s_{m+1}(i), s_{m+1}(j)] \leq r, \quad (9)$$

$$B = \sum_{i=1}^{n-m} \sum_{j=1, j \neq i}^{n-m} \#d[s_m(i), s_m(j)] \leq r, \quad (10)$$

where A is the counts that two subsequences are similar with length m , and B is the counts that two subsequences are similar with length $m+1$. By constraining $j \neq i$, self-counting is avoided. Then sample entropy is calculated as:

$$SampEn(m, r) = -\log_2\left(\frac{A}{B}\right) \quad (11)$$

Since A is always less than or equal to B , the ratio A/B is an unbiased conditional probability less than or equal to unity. In addition to self-counting, another difference between $SampEn$ and $ApEn$ is the position of logarithm. The sum of all subsequences is inside the logarithm in $SampEn$ and outside in $ApEn$. This operation reduces the probability of undefined logarithms when self-counting is not allowed. Sample entropy demonstrates improved relative consistency in comparison to approximate entropy and provides a more effective means of quantifying regularity in a system [15].

In conclusion, entropy rate estimated by the data compression algorithm measures regularity of time series based on

the length of new subsequences. While approximate entropy and sample entropy rely on the conditional probability that close patterns for m observations remain close on the next incremental comparisons. All of them take into account the evolution of the series over time. A regular evolution leads to fixed patterns in the series, further resulting in an increment of conditional probability and a shorter length of the compressed series. In contrast, a random evolution leads to the opposite outcome.

In practice, entropy rate is commonly used to measure regularity of categorical time series, such as human mobility [9], human economic behavior [10], human online life [11], web browsing behavior [12] and patient health records [13]. Approximate entropy and sample entropy are used extensively in physiological [42], medical [43], [44], economic [19], [20], and climatic [21]–[23] applications, as well as in gait analysis [24], [25], and battery health prognostics [26], [27]. However, there has been a lack of studies that apply these entropy-based metrics to assess the regularity of human physical activities. As a result, the suitability of these entropy models in measuring the regularity of human activities remains to be verified.

D. Adapting entropy rate to samples with continuous values

To apply entropy rate to longitudinal sensor data that can reflect an individual's physical activities, we modify the estimation procedure of entropy rate. As outlined in Equation (4), calculating Λ_i is essential for estimating entropy rate. In its original definition, Λ_i represents the length of the shortest subsequence which starts from position i and never exists previously. While this definition is appropriate for categorical time series, it is too rigid for time series with continuous values, such as step counts or calorie expenditures, where small numerical differences can still be considered as equivalent status. Therefore, we generalize the definition of Λ_i by replacing existence with similarity. Based on distance function $d[s_m(i), s_m(j)]$ in equation (5), we consider two subsequences are similar if $d[s_m(i), s_m(j)] \leq r$. Λ_i is redefined as the length of the shortest subsequence which starts from position i and never exists similar subsequences previously. This modification not only enables the entropy rate to be applied to continuous series, but also makes the comparison among entropy rate, approximate entropy, and sample entropy in a more fair manner, due to the same parameter r .

IV. VALIDATION OF ENTROPY MODELS

In this part, we validate the applicability of entropy rate, approximate entropy, and sample entropy for measuring the regularity of human physical activities using longitudinal sensor data.

A. Artificial physical activity data

To validate the applicability of entropy rate, approximate entropy, and sample entropy, we need to know the real regularity of longitudinal sensor data, or at least the relative regularity among these data, as the ground truth. However, the

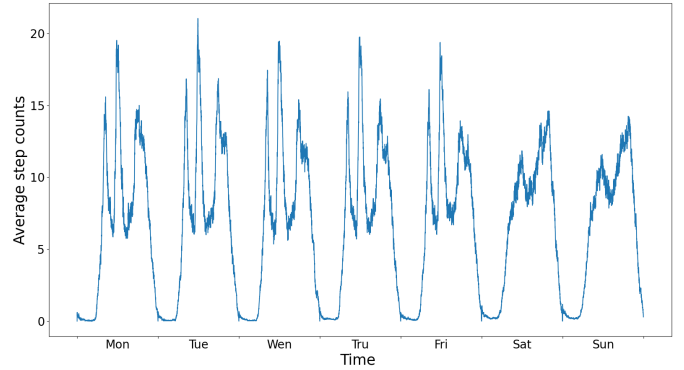


Fig. 1. Average step counts per minute within a week across users from a real-world dataset.

diversity of individual lifestyles results in a mixture of varying kinds of regularity and a multitude of random noises in real-world sensor samples. This makes it challenging to manually distinguish which samples exhibit a higher degree of regularity.

Despite the complexity of real-world data, it is possible to construct artificial physical activity data with controllable randomness to obtain a relative regularity. Moreover, the artificial data should be considered as a simulation of activity patterns that occur in real life. An analysis of a real-world dataset is depicted in Figure 1, which reveals the average step counts per minute within a week across hundreds of users. Similar shapes from Monday to Friday indicate distinct circadian rhythms existed in human physical activities. Three prominent peaks can be identified in the morning, noon, and evening of weekdays, which indicates collective exercise preference. Based on these observations, a basic activity pattern was constructed for the artificial data, that is, exercise at preference time every day. The inherent randomness in human life, however, affects the occurrence time, duration, and intensity of exercise from day to day. Additionally, trivial activities in daily life, such as housework, may also occur randomly. By controlling the degree of randomness in these elements, artificial physical activity data can be generated with known relative regularity.

More specifically, we model artificial physical activity data as the superposition of two types of physical activities, which are exercise and trivial activity. Exercise usually lasts for a long period of time with a steady intensity, and is accompanied by some preference in terms of timing. The trivial activity, by contrast, only lasts a few minutes, fluctuates in intensity, and occurs randomly throughout the day. Mathematically, artificial data of one day can be expressed as

$$Y = \sum_{i=1}^M E_i(t_i, d_i, int_i) + \sum_{j=1}^N TA_j, \quad (12)$$

where E_i denotes i -th exercise of the day, and its parameters, t_i , d_i , and int_i represent occurrence time, duration, and intensity of the exercise, respectively. TA_j is the j -th trivial activity of the day. Trivial activities can be considered as noise in daily life and are assumed to be independently and identically distributed. A complete artificial sample consists of data from several consecutive days. For simplicity, we assume that data from different days are independent.

Initially, we consider a completely regular scenario where exercise of constant intensity and equal duration are performed at the same time every day. Taking the minute-level step counts data as an example, three exercises that occur every day at 8:00, 12:00, and 20:00 with an intensity of 100 steps per minute and a duration of 60 minutes constitute a very regular sample. In this case, M is set to 3, and the parameters of exercise are also fixed on daily basis. To simulate the variability of exercise, we add normally distributed disturbances to the occurrence time, duration, and intensity of exercise. These disturbances are denoted as $dis_t \sim N(0, \sigma_t)$, $dis_d \sim N(0, \sigma_d)$, $dis_{int} \sim N(0, \sigma_{int})$, respectively. The standard deviation determines the degree of disturbance. A smaller standard deviation results in a more regular exercise pattern, while a larger standard deviation increases the level of disruption to the regularity. For trivial activities, their duration are generated from a geometric distribution with a mean of 3 minutes, and they are randomly placed throughout the day. The step counts per minute in trivial activities are sampled from a uniform distribution range from 20 to 150. Due to the random nature of trivial activity, an increase in the number of trivial activities leads to a decline in regularity. Therefore, the number of trivial activities, N , can be also used to control the degree of disturbance, like standard deviations. In Fig.2, we visualize four artificial samples, and each sample contains only one specific disturbance. The title of subplot indicates the type and degree of the disturbance. The x-axis is the time of day, and each line in the subgraph is the step counts waveform of a day. The y-axis is the number of days.

Using different parameters, we construct two-week artificial samples under these disturbances. The standard deviation of occurrence time ranges from 0 to 180 minutes in 30-minute increments and the standard deviation of duration ranges from 0 to 60 minutes in 10-minute increments. The standard deviation for intensity, measured in terms of step counts per minute, is varied from 0 to 100 in increments of 10. The disturbances from normal distributions are limited to $\pm\sigma$ to avoid cases of overlapping of exercise or meaningless values such as negative duration or step counts. The number of trivial activities per day ranges from 0 to 50. For each parameter of disturbances, we construct 100 samples. Fig.3 shows the entropy rate, approximate entropy, and sample entropy of artificial samples with different parameters. The noise filter r equals 10 to distinguish the smallest disturbance and the length of template m is 2, as [42] suggested. The solid lines represent trends of average value, and the shaded parts are 95% confidence intervals.

From Fig.3a and Fig.3b, we can find that approximate entropy and sample entropy remain unchanged with different disturbances on exercise's occurrence time and duration. Although the duration and occurrence time of exercise varies, the conditional probability that similar subsequences of length m stay consistent at the next position is almost invariable. This stability is a result of the reliance of approximate entropy and sample entropy on the evolution of small-scale subsequences. Variations in the duration and occurrence time of exercise can be considered as alterations in macro regularity, and they are difficult to distinguish at a small-scale. In the analysis

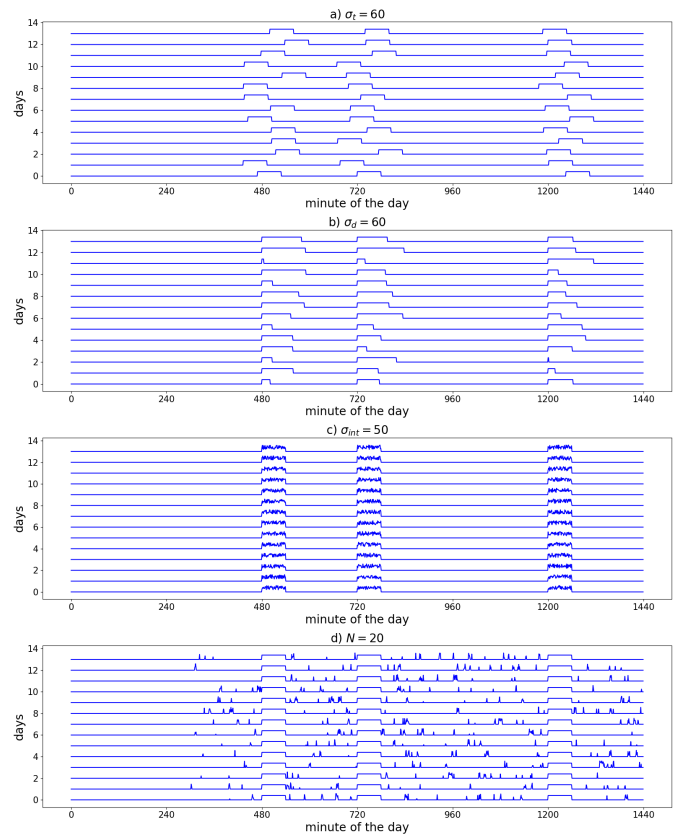


Fig. 2. Visualization of artificial physical activity samples with four types of disturbance.

of approximate entropy and sample entropy, the length of template m determines the size of subsequences. Due to the small value of m , approximate entropy and sample entropy are less sensitive to such macro irregularities. While the entropy rate increases with the degree of disturbance. The entropy rate increases greatly from zero disturbance to a small disturbance, then the increment of entropy rate becomes smaller with the increment of disturbance. We can understand this phenomenon from the perspective of lossless data compression. For the completely regular sample, all activities are exactly the same and can be compressed as one symbol. However, a small disturbance leads to minor variations among activities, resulting in the need to compress each unique activity as a separate symbol. As the disturbance level increases, the number of identical activities decreases and all activities must be compressed as distinct symbols, leading to a gradual decrease in the growth rate of entropy rate. Fig.3c shows the performance of three entropy models on artificial samples under different degree of disturbance on exercise's intensity. The disturbance can be regarded as noise of varying magnitudes applied to the original signal. It is observed that the approximate entropy increases as the magnitude of the noise increases. The larger the noise, the wider range of values at each time points. And the probability that similar subsequences are still close at the next moment becomes smaller. Entropy rate is also increasing and leveling off faster. This trend is similar to Fig.3a and Fig.3b for the same reason. We zoom in on the sample entropy curve in

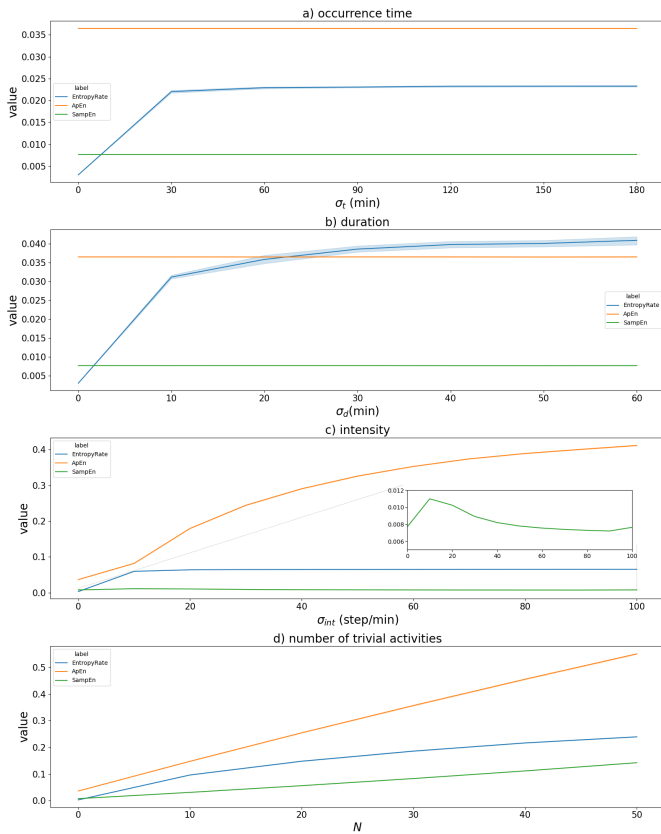


Fig. 3. Entropy rate, approximate entropy, and sample entropy of artificial samples under different degree of disturbance.

Fig.3c, which increases and then decreases as disturbance increases. This is because sample entropy sums over the numerator and denominator of all subsequences separately when calculating the conditional probability. As shown in Fig.3d, all three entropy models increase with the increment of trivial activities, which means all of them are able to distinguish the amount of noise in physical activity data.

Summarizing the performance of three kinds of entropy on artificial physical activity data, we can find that approximate entropy and sample entropy can not distinguish the macroscopic variation of physical activity, such as duration and occurrence time, as the entropy rate can. Additionally, approximate entropy and entropy rate are able to identify the magnitude and amount of noise in physical activity data, while sample entropy can only identify the amount of noise.

B. Real-world physical activity data

The artificial physical activity data can be used to validate the applicability of entropy rate, approximate entropy, and sample entropy under specific situations. However, it is not sufficient to evaluate the applicability of these three kinds of entropy based only on the results of artificial data, given that the real situation is more complex than artificial data. A demonstration of how well these three types of entropy perform on real-world physical activity data will be more convincing. Although it is hard to know the real regularity of real-world samples, predictability can serve as a proxy for

regularity since regular data is more predictable.

Predictability is a measure of how well future values of a time series can be forecasted [45]. To evaluate the predictability of a time series with length n , a forecasting approach is employed starting at an initial position i . The i -th data point is predicted based on the previous data points, and this procedure is repeated until $i = n$. The predictability of the time series is determined by the average error of the predictions made at each position. It is important to note that the predictability of a time series is not only dependent on the degree of regularity present in the series, but also on the choice of forecasting models. To mitigate the effect of model choice on predictability, we employ four widely used time series forecasting models, including both classical statistical models and deep learning models, as follows.

- 1) Exponential smoothing (ES). Forecasts produced using ES are weighted averages of past observations, with the weights decaying exponentially as the observations get older [46]. ES method can capture the trend and seasonality of time series by applying exponential smoothing recursively [47].
- 2) Prophet. Prophet forecasts time series based on an additive model where non-linear trends are fit with yearly, weekly, and daily seasonality, plus holiday effects [48].
- 3) Recurrent neural networks (RNN). RNN is a class of neural networks allowing output from nodes affects subsequent input, which is able to capture long-term temporal dependencies. Long Short-Term Memory (LSTM) [49] and Gated Recurrent Unit (GRU) [50] are two popular variants of RNN, which have been shown to achieve state-of-the-art results in applications with time series.
- 4) N-BEATS. N-BEATS is a deep neural architecture based on backward and forward residual links and a very deep stack of fully-connected layers [51]. The residual connection enables the stacks to focus on predicting errors from the previous stack, which implements an automatic time series decomposition.

We first validate the performance of these forecasting models in artificial physical activity data. To reduce computational overhead, the data is first reduced from minute-level to hour-level by summing the values per minute within one hour. In this way, the total length of 30-day artificial samples is reduced to 720. We historically forecast the entire sample starting from the first 1/3 time steps, so that the forecast horizon is long enough and sufficient historical data is available for the initial forecast. The mean absolute error (MAE) is used as a metric to evaluate the forecasting performance of the models. Since the impact of disturbance on intensity will cancel out after the summation process, we validate the performance of forecasting models in hourly artificial samples with the presence of other three kinds of disturbances. The results, as shown in Fig. 4, demonstrate that the MAE of the models increases with the increase of the disturbance level. The trend of prediction error is consistent with our expectation of regularity, and this indicates that the prediction error can serve as a reliable proxy for regularity. The correlation coefficient between the

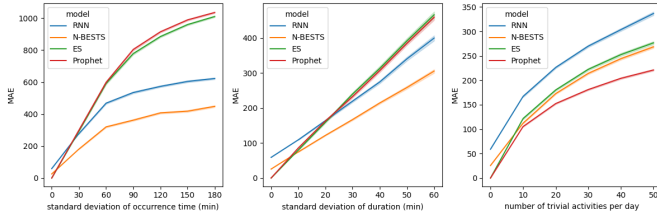


Fig. 4. Performance of forecast models on hourly artificial physical activity data.

TABLE II
MAE OF FORECASTING MODELS ON REAL-WORLD DATASETS.

Models	SJTU (step counts)	Fitbit (step counts)	Fitbit (calorie)	Fitbit (intensity)
ES	428 ± 148	298 ± 137	24.4 ± 12.6	9.61 ± 3.82
Prophet	466 ± 160	323 ± 146	28.1 ± 13.1	10.87 ± 3.89
RNN	346 ± 124	266 ± 134	23.1 ± 11.9	9.05 ± 3.69
N-BEATS	353 ± 123	294 ± 144	24.8 ± 12.2	9.52 ± 3.62

regularity measurements and the prediction error can then be used to evaluate the performance of these metrics on real-world physical activity data.

We then present two real-world physical activity datasets. (1) SJTU dataset¹. This dataset contains minute-level step counts data of 686 users over 30 consecutive days, collected through a smartphone application called SJTU Health with the consent of users. (2) Fitbit dataset². This dataset is generated from thirty eligible Fitbit users who consented to submit approximately one month of personal tracker data, including minute-level step counts, calorie expenditures, and exercise intensities. The values of intensity range from 0 to 3, representing sedentary, lightly active, moderately active, and very active respectively. The intensity classification is determined by proprietary algorithms from Fitbit.

We historically forecast these real-world samples from different users after scaling them to hourly data, just as we forecast artificial samples. The results, as presented in Table II, demonstrate the mean and standard deviation of the MAE of each forecasting model on real-world datasets. Among all models, RNN exhibits the lowest average prediction error on all datasets. The second-best performance is achieved by the N-BESTS. These deep learning-based models perform better than traditional models, like ES and Prophet.

Before calculating entropy rate, approximate entropy, and sample entropy for these real-world samples, the noise filter r needs to be assigned. The value of r should be chosen carefully, as a small r can result in a high sensitivity to noise, leading to the masking of real regularities by tiny fluctuations. On the other hand, if r is set too large, it will reduce the ability to distinguish noise and result in all entropy values being the same in an extreme scenario where r approaches infinity. Here, we assign r equal to the standard deviation of all data points in the dataset, considering standard deviation can reflect the amount of dispersion of data. In Table III, we list the mean and

TABLE III
MEAN AND STANDARD DEVIATION OF ALL DATA POINTS IN REAL-WORLD DATASETS.

Datasets	SJTU (step counts)	Fitbit (step counts)	Fitbit (calorie)	Fitbit (intensity)
Mean	402	320	98	12
Std	895	690	62	21

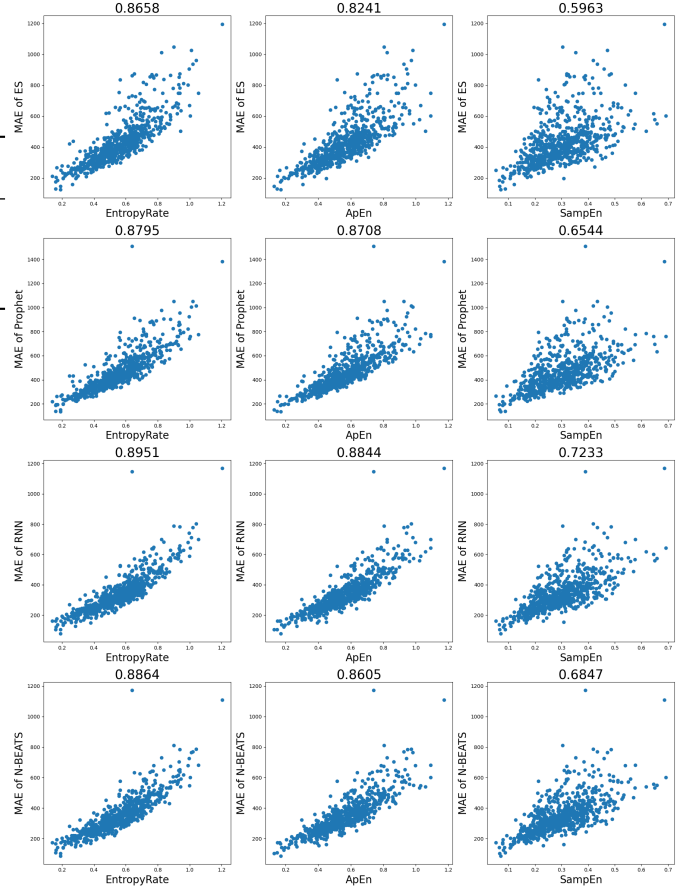


Fig. 5. Scatter plot of MAE and entropy on SJTU dataset.

standard deviation in SJTU and Fitbit datasets after scaling. For approximate entropy and sample entropy, the length of template m is 2, as [42] suggested.

After obtaining prediction errors and three kinds of entropy for all real-world samples, we calculate the Spearman correlation coefficient between prediction error and entropy to determine the best regularity measurement. The Spearman correlation coefficient is a statistical measure of the strength of a monotonic relationship between paired data. Since regular data can be predicted better, samples with smaller prediction errors should have smaller entropy. Among three kinds of entropy, the entropy with the highest Spearman correlation coefficient is considered as the best regularity measurement.

In Fig.5, we display the scatter plots of prediction error versus three kinds of entropy for four forecasting models in the SJTU dataset. Each subplot is titled with the Spearman correlation coefficient between prediction error and entropy. The results show that for all forecasting models, the correlation

¹Data is available under reasonable requests.

²<https://www.kaggle.com/datasets/arashnic/fitbit>

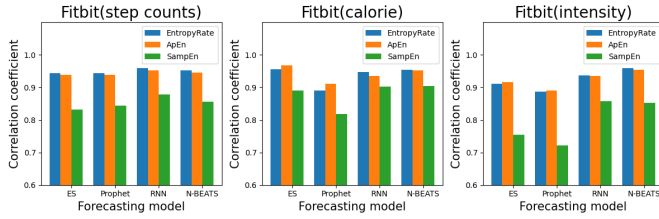


Fig. 6. Correlation coefficients between prediction error and entropy on Fitbit dataset.

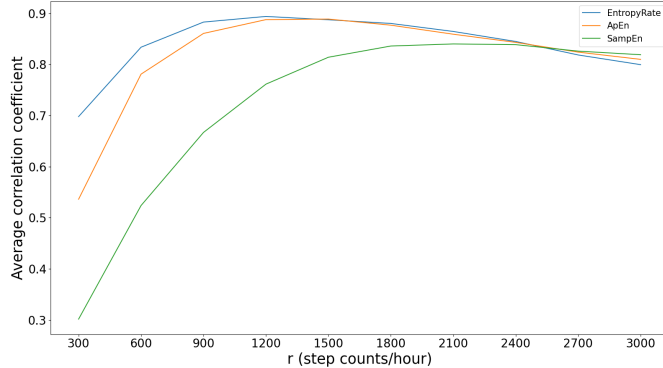


Fig. 7. Average correlation coefficient between prediction errors and entropy under varying noise filters in the SJTU dataset.

between entropy rate and prediction error is stronger compared to that of approximate entropy and sample entropy. For all kinds of entropy, the correlation coefficient between prediction error of the best-performing RNN model and entropy is greater than other models. The correlation coefficient between entropy rate and error of RNN is the highest, which is 0.8951. Fig.6 presents correlation coefficients between entropy and prediction error on the Fitbit dataset. It is observed that the correlation coefficients between sample entropy and prediction error are notably lower than those between entropy rate and approximate entropy. The performance of approximate entropy is found to be comparable to that of entropy rate. And the majority of their correlation coefficients are above 0.9.

We also perform a sensitivity analysis to investigate the impact of different noise filters r on the performance of three entropy models. Fig.7 presents the mean of correlation coefficients between prediction errors from different models and entropy under varying r in the SJTU dataset. The results reveal that the average correlation coefficients of all three entropy types initially increase and then decrease as r increases. This phenomenon is consistent with our previous analysis that entropy value calculated from both too small and too large r can not reflect the real regularity. Additionally, entropy rate displays a higher correlation coefficient than approximate entropy and sample entropy across a wide range of parameters. Especially when r is small, the advantage of entropy rate is more obvious. Moreover, with the increase of r , the correlation coefficient of entropy rate does not decline significantly. These indicate that the entropy rate is more robust to the choice of parameter r .

In conclusion, entropy rate is more relevant to the prediction error of real-world physical activity samples than approximate

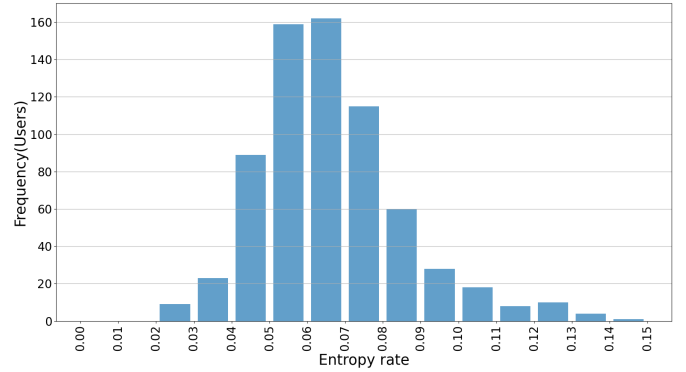


Fig. 8. Distribution of entropy rate across users in the SJTU dataset.

entropy and sample entropy. Combining their performance on artificial data, entropy rate is more suitable than approximate entropy and sample entropy to measure the regularity of human physical activity.

V. HUMAN PHYSICAL ACTIVITY REGULARITY

A. Interpersonal variation of physical activity regularity

Based on entropy rate, we quantify the regularity of physical activity for different individuals. As shown in Fig.8, we present the distribution of entropy rate of minute-level data for all users in the SJTU dataset. The average entropy rate of all users is 0.066 bits. In this context, entropy rate refers to the average amount of newly generated information for each update of the user's physical activity state. A entropy rate of 0.066 bit can be interpreted as the user's physical activity state of next minute could be found on average in any of $2^{0.066} \approx 1.046$ states, which also means the user's physical activity state is determinable most of the time. These results are reasonable given that sedentary and restful activities typically occupy a significant portion of the day for most individuals. The SJTU dataset reveals that users on average have 1280 minutes per day with 0 step counts. Additionally, the ordered structure of the data, such as prolonged periods of a single state or regular alternations between states, further reduces the number of potential states for the next moment. Fig.8 also indicates distinct variations exist among individuals' regularity. Such interpersonal variability on regularity reflects different lifestyle preferences of people.

Fig.9 visualizes step counts data for two specific users with different entropy rates over the month-long observation period. User 1 displays a clear exercise pattern, characterized by consistent daily walks at set times in the morning, noon, and evening. Conversely, the data for user 2 exhibits a more unpredictable and chaotic pattern, with various fluctuating activities that occur at irregular times. As a result, the entropy rate for user 2 (0.1015 bit) is significantly higher than that for user 1 (0.0411 bit).

We also explore factors that may contribute to variations in regularity among individuals. To this end, we derive two features from the original physical activity data for all users. The first feature is the daily step counts, which serves as a measure of the total amount of physical activity. The second

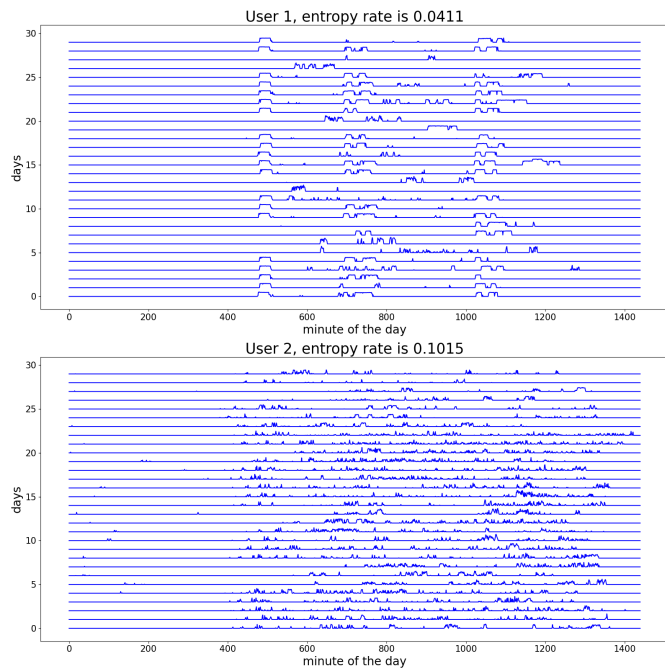


Fig. 9. Case study of two specific users with different entropy rate over the 30 days

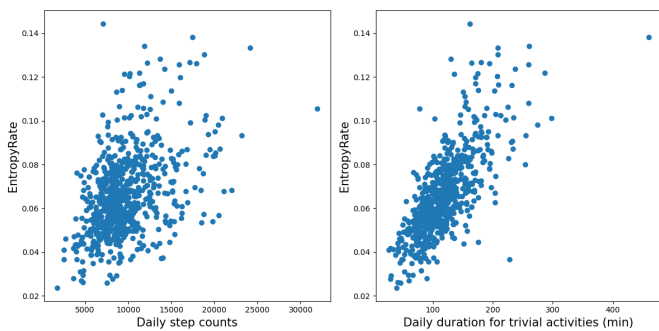


Fig. 10. Scatter plot of regularity and characteristics of physical activities across all users in SJTU dataset.

feature is the daily duration of trivial activities, which can be used as a proxy for the composition of physical activity. The trivial activity is determined based on duration and intensity. Physical activities that last less than 10 minutes or 1000 step counts are classified as trivial activity, because activities with more than 1000 step counts in 10 minutes are generally regarded as effective exercise [52]. Fig.10 displays the relationship between entropy rate and these two features for all users in SJTU dataset. When daily step counts is low, the entropy rate tends to be low as well, as there is limited physical activity. However, as the daily step count increases, the entropy rate exhibits a wider range of regularity. In general, daily step counts is weakly correlated with entropy rate, and the correlation coefficient is 0.4403. On the other hand, the average duration spent on trivial activities per day shows a great positive correlation with entropy rate, and the correlation coefficient is 0.7290. Trivial activities tend to exhibit more randomness in both the occurrence time and intensity compared to exercise, which leads to a decrease in regularity

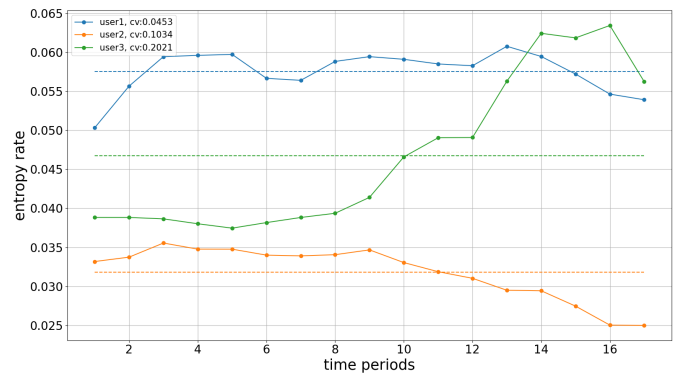


Fig. 11. Illustration of entropy rate sequences from three users. The dotted line represents the mean value of sequence.

as the time spent on trivial activities increases. Overall, the results suggest that the regularity of human physical activity is not determined by the amount of activity, but rather by the composition of activities.

B. Intrapersonal variation of physical activity regularity

Intrapersonal variation of regularity describes temporal variability of the same individual's physical activity habits. For each user in SJTU dataset, we recursively slide through the entire sequence in a 14-day window (e.g. day1-day14, day2-day15, ..., day17-day30). In this way, we divide user's original time series of 30 days into 17 two-week time periods. The entropy rate of consecutive time periods forms an entropy rate sequence, which indicates how an individual's regularity changes over time. Fig.11 illustrates entropy rate sequences from three users with different degrees of variability. We utilize Coefficient of Variation (CV) to quantify the degree of variation, which is the ratio of the standard deviation to the mean. The entropy rate sequence of user 1 has the smallest CV value, 0.04, among these three users, and it fluctuates around the blue dashed line, which is the mean of entropy rate sequence. This indicates the physical activity regularity of user1 is almost constant. The entropy rate sequence of user 2 decreases slightly in the last several time periods, and it has a CV value of 0.10. For user 3, the entropy rate increases significantly over time in the second half. Therefore, it has the highest CV value of 0.20.

We also calculated the coefficient of variation of entropy rate sequence for all users, and plot its cumulative distribution function (CDF) in Fig.12. We can find that about 40% of users have CV values less than 0.05, and more than 80% of the users have CV values less than 0.1. This indicates the majority of people have stable physical activity habits.

VI. CONCLUSION

Measuring regularity of physical activity is important for understanding an individual's overall physical activity behavior and for developing effective interventions to promote physical activity and improve health. In this study, we prove that entropy rate can be regarded as an appropriate measurement for regularity of human physical activities. On artificial physical activity data with controllable randomness, entropy rate

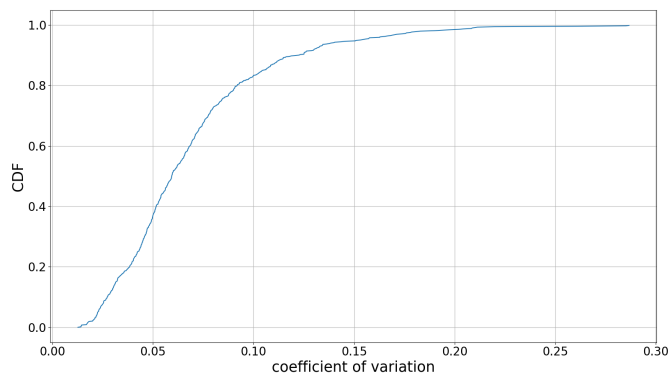


Fig. 12. Cumulative distribution function of the coefficient of variation for all users in the SJTU dataset.

exhibits the ability to identify not only the magnitude and amount of noise but also macroscopic variations of physical activities, such as duration and occurrence time. On real-world physical activity datasets, entropy rate is closely tied to the predictability of samples. The strong correlations between entropy rate and prediction errors from various forecasting models demonstrate its applicability in measuring human physical activity regularity.

As an accurate measure of regularity, entropy rate can provide us with a better understanding of patterns of physical activities on individual and population-level. Based on entropy rate, we investigate the interpersonal and intrapersonal variations of regularity for numerous users. The results show that the interpersonal variability in the regularity of physical activity engagement is readily apparent in the population. We further find that the heterogeneity in regularity among individuals is attributable to the composition of physical activities than the amount of physical activities. Individuals who spent more time on trivial activities usually tend to exhibit greater irregularity. Simultaneously, the regularity of physical activities displays negligible intrapersonal variation for the majority of individuals. The temporal stability of regularity suggests the prevalence of stable physical activity habits among the populace. In addition to the aforementioned investigations, regularity measured by entropy rate can also be advantageous in identifying subgroups necessitating intervention to foster habitual physical activity. Better targeting of interventions then can be implemented to improve health of subgroups. Furthermore, regularity can function as crucial prior knowledge in individual-level physical activity modeling and prediction.

One limitation of this study, however, is that the data recorded by wearable or other smart devices have their inherent limits. Some activities may fail to be recorded when the device is not carried along by its owner. If that happens, the regularity of the user behavior may appear to be different than it really is. In this sense, it might be wise to treat the non-zero data points differently from zero ones, such that "real" activities may be prioritized in the analysis. It would also be of interest to develop a multifaceted approach to assessing the regularity of human activities, leveraging diverse data sources such as activity data, weather data, and exter-

nal incentives. Such an integrated approach could provide a more comprehensive and nuanced understanding of the factors shaping physical activity patterns and inform the development of targeted interventions to promote regular physical activity engagement.

REFERENCES

- [1] World Health Organization, *WHO guidelines on physical activity and sedentary behaviour*, 2020.
- [2] Ž. Pedišić, "Measurement issues and poor adjustments for physical activity and sleep undermine sedentary behaviour research—the focus should shift to the balance between sleep, sedentary behaviour, standing and activity," *Kinesiology*, vol. 46, no. 1, pp. 135–146, 2014.
- [3] T. K. Monk *et al.*, "The Social Rhythm Metric: An instrument to quantify the daily rhythms of life," *Journal of Nervous and Mental Disease*, 1990.
- [4] T. H. Monk *et al.*, "A simple way to measure daily lifestyle regularity," *Journal of sleep research*, vol. 11, no. 3, pp. 183–190, 2002.
- [5] C. Tudor-Locke, J. E. Williams and J. P. Reis "Utility of pedometers for assessing physical activity," *Sports medicine*, vol. 32, no. 11, pp. 795–808, 2002.
- [6] G. Goulet-Langlois *et al.*, "Measuring regularity of individual travel patterns," *IEEE Transactions on Intelligent Transportation Systems*, vol. 19, no. 5, pp. 1583–1592, 2017.
- [7] T. M. Cover and J. A. Thomas, "Entropy, Relative Entropy, and Mutual Information," *Elements of information theory*, 2nd ed., JOHN WILEY & SONS, 2006, pp. 13–16.
- [8] A. Delgado-Bonal and A. Marshak, "Approximate entropy and sample entropy: A comprehensive tutorial," *Entropy*, vol. 21, no. 6, pp. 541–578, 2019.
- [9] C. Song *et al.*, "Limits of predictability in human mobility," *Science*, vol. 327, no. 5968, pp. 1018–1021, 2010.
- [10] C. Krumme *et al.*, "The predictability of consumer visitation patterns," *Scientific reports*, vol. 3, no. 1, pp. 1–5, 2013.
- [11] R. Sinatra and M. Szell, "Entropy and the predictability of online life," *Entropy*, vol. 16, no. 1, pp. 543–556, 2014.
- [12] J. Kulshrestha *et al.*, "Web routiness and limits of predictability: Investigating demographic and behavioral differences using web tracking data," in *Proceedings of the International AAAI Conference on Web and Social Media*, 2021, pp. 327–338.
- [13] D. Dahlem, D. Maniloff, and C. Ratti, "Predictability bounds of electronic health records," *Scientific reports*, vol. 5, no. 1, pp. 1–9, 2015.
- [14] S. M. Pincus, "Approximate entropy as a measure of system complexity," *Proceedings of the National Academy of Sciences*, vol. 88, no. 6, pp. 2297–2301, 1991.
- [15] J. S. Richman and J. R. Moorman, "Physiological time-series analysis using approximate entropy and sample entropy," *American Journal of Physiology-Heart and Circulatory Physiology*, vol. 278, no. 6, pp. H2039–H2049, 2000.
- [16] D. E. Lake, J. S. Richman, M. P. Griffin, and J. R. Moorman, "Sample entropy analysis of neonatal heart rate variability," *American Journal of Physiology-Regulatory, Integrative and Comparative Physiology*, 2002.
- [17] U. R. Acharya *et al.*, "Automated diagnosis of epileptic EEG using entropies," *Biomedical Signal Processing and Control*, vol. 7, no. 4, pp. 401–408, 2012.
- [18] W. Chen, J. Zhuang, W. Yu, and Z. Wang, "Measuring complexity using fuzzyen, apen, and sampen," *Medical engineering & physics*, vol. 31, no. 1, pp. 61–68, 2009.
- [19] A. Delgado-Bonal, "Quantifying the randomness of the stock markets," *Scientific reports*, vol. 9, no. 1, pp. 1–11, 2019.
- [20] M. Xu, P. Shang, and J. Huang, "Modified generalized sample entropy and surrogate data analysis for stock markets," *Communications in Nonlinear Science and Numerical Simulation*, vol. 35, pp. 17–24, 2016.
- [21] J. Meng *et al.*, "Complexity-based approach for El Niño magnitude forecasting before the spring predictability barrier," *Proceedings of the National Academy of Sciences*, vol. 117, no. 1, pp. 177–183, 2020.
- [22] L. Zhang *et al.*, "Identification and application of the most suitable entropy model for precipitation complexity measurement," *Atmospheric Research*, vol. 221, pp. 88–97, 2019.
- [23] S. F. A. Xavier *et al.*, "An application of sample entropy to precipitation in Paraíba State, Brazil," *Theoretical and Applied Climatology*, vol. 136, no. 1, pp. 429–440, 2019.
- [24] S. Ramdani *et al.*, "On the use of sample entropy to analyze human postural sway data," *Medical engineering & physics*, vol. 31, no. 8, pp. 1023–1031, 2009.

- [25] J. D. McCamley *et al.*, "On the calculation of sample entropy using continuous and discrete human gait data," *Entropy*, vol. 20, no. 10, pp. 764–784, 2018.
- [26] X. Hu, S. E. Li, Z. Jia, and B. Egardt, "Enhanced sample entropy-based health management of Li-ion battery for electrified vehicles," *Entropy*, vol. 64, no. 1, pp. 953–960, 2014.
- [27] A. Widodo, M. Shim, W. Caesarendra, and B. Yang, "Intelligent prognostics for battery health monitoring based on sample entropy," *Expert Systems with Applications*, vol. 38, no. 9, pp. 11763–11769, 2011.
- [28] A. Zisberg, N. Gur-Yaish, and T. Shochat, "Contribution of routine to sleep quality in community elderly," *Sleep*, vol. 33, no. 4, pp. 509–514, 2010.
- [29] M. S. Câmara *et al.*, "Lifestyle regularity measured by the social rhythm metric in Parkinson's disease," *Chronobiology international*, vol. 22, no. 5, pp. 917–924, 2005.
- [30] T. F. Campos *et al.*, "Regularity of daily activities in stroke," *Chronobiology international*, vol. 25, no. 4, pp. 611–624, 2008.
- [31] S. Saeb *et al.*, "Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: an exploratory study," *Journal of medical Internet research*, vol. 17, no. 7, pp. e4273, 2015.
- [32] W. Wang *et al.*, "Sensing behavioral change over time: Using within-person variability features from mobile sensing to predict personality traits," *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, vol. 2, no. 3, pp. 1–21, 2018.
- [33] A. J. Phillips *et al.*, "Irregular sleep/wake patterns are associated with poorer academic performance and delayed circadian and sleep/wake timing," *Scientific reports*, vol. 7, no. 1, pp. 1–13, 2017.
- [34] C. Shang *et al.*, "Implicit irregularity detection using unsupervised learning on daily behaviors," *IEEE Journal of Biomedical and Health Informatics*, vol. 24, no. 1, pp. 131–143, 2019.
- [35] M. Marschollek, "Physical activity event regularity and health outcome—'Undiscovered country' in cohort accelerometer data," *Digital Healthcare Empowering Europeans*, pp. 657–659, 2015.
- [36] D. Fischer, E. B. Klerman, and A. J. Phillips, "Measuring sleep regularity: theoretical properties and practical usage of existing metrics," *Sleep*, vol. 44, no. 10, pp. zsab103, 2021.
- [37] w. Witting *et al.*, "Alterations in the circadian rest-activity rhythm in aging and Alzheimer's disease," *Biological psychiatry*, vol. 27, no. 6, pp.563–572, 1990.
- [38] A. Lempel and J. Ziv, "On the complexity of finite sequences," *IEEE Transactions on information theory*, vol. 2, no. 1, pp. 75–81, 1976.
- [39] I. Kontoyiannis, P. H. Algoet, Y. M. Suhov, and A. J. Wyner, "Nonparametric entropy estimation for stationary processes and random fields, with applications to English text," *IEEE Transactions on information theory*, vol. 33, no. 3, pp. 1319–1327, 1998.
- [40] S. M. Pincus, "Approximate entropy (ApEn) as a complexity measure," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 5, no. 1, pp. 110–117, 1995.
- [41] S. M. Pincus, "Approximating Markov chains," *Proceedings of the National Academy of Sciences*, vol. 89, no. 10, pp. 4432–4436, 1992.
- [42] S. M. Pincus and A. L. Goldberger, "Physiological time-series analysis: what does regularity quantify?" *American Journal of Physiology-Heart and Circulatory Physiology*, vol. 266, no. 4, pp. H1643–H1656, 1994.
- [43] X. Zhang *et al.*, "EMG-torque relation in chronic stroke: a novel EMG complexity representation with a linear electrode array," *IEEE Journal of Biomedical and Health Informatics*, vol. 21, no. 6, pp. 1562–1572, 2016.
- [44] A. Lanata *et al.*, "Complexity index from a personalized wearable monitoring system for assessing remission in mental health," *IEEE Journal of Biomedical and health Informatics*, vol. 19, no. 1, pp. 132–139, 2014.
- [45] R. K. Guntu *et al.*, "Wavelet entropy-based evaluation of intrinsic predictability of time series," *Chaos: An Interdisciplinary Journal of Nonlinear Science*, vol. 30, no. 3, pp. 033117, 2020.
- [46] R. J. Hyndman and G. Athanasopoulos, "Exponential smoothing," *Forecasting: principles and practice*, OTexts, 2018, pp. 238.
- [47] P. R. Winters, "Forecasting sales by exponentially weighted moving averages," *Management science*, vol. 6, no. 3, pp. 324–342, 1960.
- [48] S. J. Taylor and B. Letham, "Forecasting at scale," *The American Statistician*, vol. 72, no. 1, pp. 37–45, 2018.
- [49] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural computation*, vol. 9, no. 8, pp. 1735–1780, 1997.
- [50] K. Cho *et al.*, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," *arXiv preprint arXiv:1406.1078*, 2014.
- [51] B. N. Oreshkin, D. Carпов, N. Chapados, and Y. Bengio, "N-BEATS: Neural basis expansion analysis for interpretable time series forecasting," *arXiv preprint arXiv:1905.10437*, 2019.
- [52] S. J. Marshall *et al.*, "Translating physical activity recommendations into a pedometer-based step goal: 3000 steps in 30 minutes," *American journal of preventive medicine*, vol. 36, no. 5, pp.410–415, 2009.