time-series can be characterized by its power spectrum density. The power spectrum density shows how the periodic frequencies (e.g. seconds, minutes, hours, days, and years) of the time-series are distributed over an entire frequency range, (Hipel and McLeod 1994). If the power spectrum density of a given time-series is distributed uniformly over the entire frequency range, the pattern of such time-series moves randomly up and down as times goes, like stock price.

Any time series can be expressed as a combination of cosine and sine waves with differing periods and amplitudes (maximum/minimum value during the cycle). This can be used to examine the periodic behavior in a time series. The normalized cumulated periodogram (described hereafter as NCP) is one of the common methods used for identifying periodicity of a given time-series in frequency domain (Hipel and McLeod 1994).

For a given n stationary time series ($y_1,...,y_n$), the periodogram function ($I(f_j)$), which shows spectral density of the time series at each frequency, is calculated as shown in Equation (3) (Hipel and McLeod 1994).

$$I(f_j) = \frac{2}{n}\left|\sum_{l=1}^{n} y_l \exp(-2\pi i f_j l)\right| = \frac{2}{n}\left[\left(\sum_{l=1}^{n} y_l \cos(2\pi f_j l)\right)^2 + \left(\sum_{l=1}^{n} y_l \sin(2\pi f_j l)\right)^2\right]^{1/2} \quad (3)$$

where $f_j = j/n$ is the j-th frequency j=1, 2,… N', N'=[n/2] (take integer portion of n/2), |·| denotes the magnitude and $i = \sqrt{-1}$. In essence, $I(f_j)$ measures the strength (or spectral density) of the relationship between the data sequence y_n and a sinusoid with frequency f_j where 0< f_j ≤0.5 (Hipel and McLeod 1994). The NCP is defined as Equation (4).

$$C(f_k) = \frac{\sum_{j=1}^{k} I(f_j)}{n c_0^{\,2}} \quad (4)$$

where $C(\cdot)$ is the NCP which is function of the frequency f, $c_0^{\,2}$ is the estimated variance defined in Equation (4). Randomness of w_k can be tested in the time domain (by auto-correlation function) and the frequency domain (by NCP). The normalized cumulative periodogram identifies hidden periodicity of the time-series using its power spectrum density. In this paper, only the result of the NCP was described here.

If a given time series follows random walk, power spectrum density of w_k (Equation 2) shall be evenly distributed over the frequency. Therefore the NCP of w_k shows a plot of the cumulative periodogram that is a straight line joining (0, 0), and (0.5, 1) as shown in Figure 1 (the NCP has a range from 0 to 1). When testing for white noise using the NCP, the confidence limits can be drawn in parallel to the line from (0,0) to (0.5,1) on x-y plane. The limits are drawn at vertical distances $\pm K_\varepsilon / \sqrt{[(n-1)/2]}$ above and below the theoretical white noise line where the [·] means to take only the integer portion of the number inside the brackets and K_ε is a parameter for calculating confidence limits. In general, 1.36 of K_ε is taken when drawing 95% of confidence limits (Hipel and McLeod 1994). Figure 1 shows the NCP of the sequence of 200 random numbers.

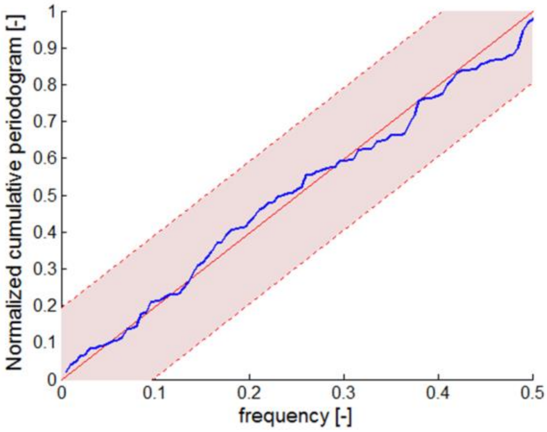

Figure 1. *Normalized cumulative periodogram of the 200 random numbers' sequence (red dotted lines mean 95% confidence limits)*

If the given time series has slowly varying periodicity the blue line shall be biased to the left (towards low frequency); otherwise, the blue line shall be biased to the right (towards high frequency). In this paper, a random walk hypothesis on an occupancy pattern (Equation 1) is tested by identifying randomness of w_k (Equation 2).

EXPERIMENTS

Occupancy pattern of two laboratories (A: mixed students in a laboratory shared by several research groups, B: building simulation laboratory) in Sunkyunkwan university campus was gathered. Arrival time and departure time of occupants were recorded using a webcam. The webcam records the photographs automatically if a motion is sensed. Based on the photographs the occupancy patterns were tabulated with time interval of 10 minutes for experiment A and time interval of 1 minute for experiment B (during Exp. B the rate of change in occupancy is relatively fast). Figure 2 shows the recorded images of two experiments.

Figure 2. *The recorded images (left: A lab, right: B lab)*

The lifestyle of the two rooms follows a "typical" university lab lifestyle; since there was no fixed/strict office hours in the labs, graduate students felt free to enter and leave the labs according to their own preferences. Figure 3 show the occupancy pattern of each experiment. In the case of the laboratories (exp. A, B), the rooms were occupied till late at night (Figure 3). The occupancy levels of each experiment fluctuate with arrival time, lunch time, and departure time. However, a closer look at the rest of the day reveals that the degree of the variations seems to be irregular. To test whether a given time-series follows a random walk behavior or not, the differenced time-series (w_k in Equation 2) was tested by the NCP.

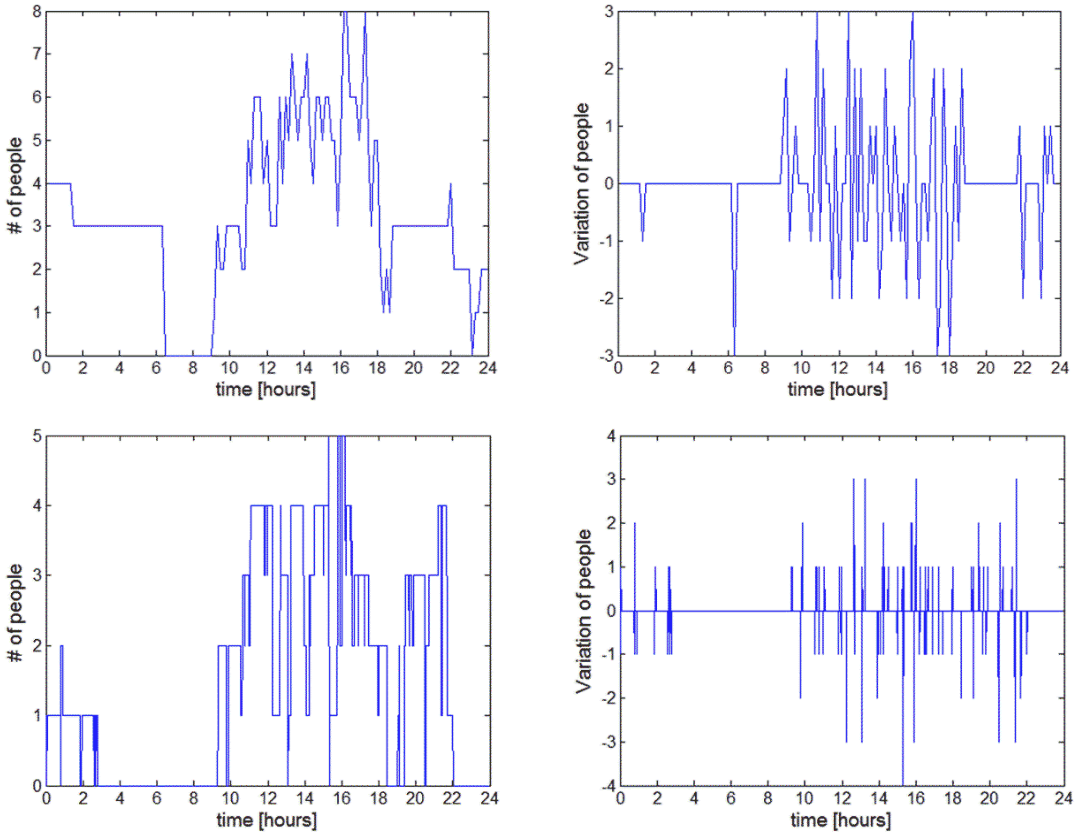

Figure 3. *Occupancy pattern (left) and differenced pattern (right) of Exp.A (top) and Exp.B (bottom)*

RESULTS

As mentioned earlier, the NCP is used for identifying a "dominant" periodicity of a given time-series in frequency domain. The NCP literally shows a "normalized cumulative sum" of the spectral density ranging from 0.0 to 0.5 of frequency. Hence the meaning of the above-mentioned "dominant" is that a strong spectral density exists between 0.0069/10 minutes - 0.0416/10 minutes frequencies and Figure 4 shows the pop-up pattern at those frequencies. Such pattern implies a certain dominant wave of cosine (or sine) exists in the given time-series.

As shown in Figure 4, the NCP of the occupancy pattern increases rapidly from (0.0069, 0.305) to (0.0138, 0.643), secondly from (0.0277, 0.685) to (0.0347, 0.728) and thirdly from (0.0347, 0.728) to (0.0416, 0.771). Each element in (0.0069, 0.305) represents the location of x, y coordinates: the number 0.0069, the first point's x-coordinate corresponds to the inverse of the sample size (1/144 = 10/24/60) and the number 0.0138, the second point's x-coordinate corresponds to 2 × 1/144, and so on.

Meanwhile, the occupant presence of Exp. A has about 4.0 - 4.8 hours and 12 hours of periodicity. Here, 4.0 - 4.8 hours of periodicity are reciprocals of the frequencies of 0.0347/10 minutes and 0.0416/10 minutes (1/0.0347/10 = 288.1 minutes = 4.8 hours, 1/0.0416/10 = 240.3 minutes = 4.0 hours). The occupancy level of typical buildings bobs up and down near the meal time (e.g. 12:00-1:00 P.M. for lunch, 6:00-7:00 P.M. for dinner) and 4 - 5 hours of periodicity seem to be fair. In addition, 12 hours of periodicity is observed with high intensity (from (0.0069, 0.305) to (0.0138, 0.643) in Figure 4). It implies that the students arrive and leave the labs at a certain point of time although there were no fixed/strict office hours in the lab during the interview with the authors.

On the other hand, the NCP of the differenced pattern (Figure 4) consistently increases from (0, 0) to (0.5, 1) within 95% confidence limits. It means that the spectral density of the given time series is evenly distributed over the frequency, and thus can be considered to be white noise. It can be inferred that the occupant presence of experiment A follows the random walk. In short, although the occupancy pattern has periodicity macroscopically, however its variations are irregular microscopically.

The result of the frequency domain analysis for experiment B is similar to that of experiment A. As shown in Figure 5, the cumulative periodogram of the original schedule considerably increases near the frequency of 0.0027-0.0035/minutes. It can be verified by checking the value of NCP, which also increases from 0.581 to 0.691 at that frequency. The frequency of 0.0035/minutes is roughly 4.8 hours (288 minutes) periodicity. As mentioned earlier, it represents a human scale of periodicity (e.g. meal time).

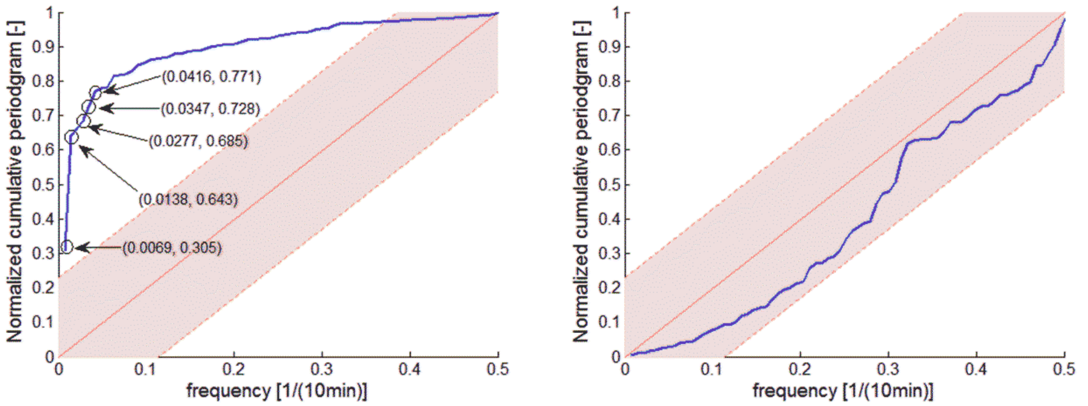

Figure 4. *Normalized cumulative periodogram of experiment A of occupancy pattern (left), and differenced pattern (right)*

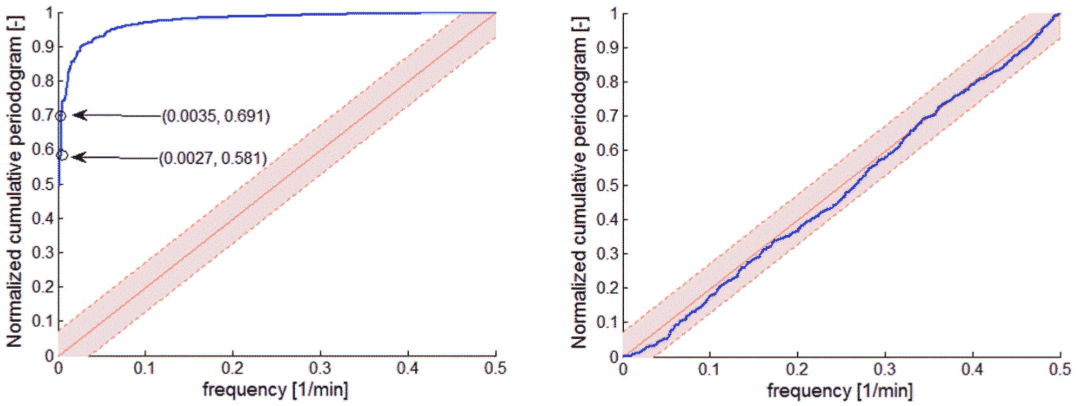

Figure 5. *Normalized cumulative periodogram of experiment B of occupancy pattern (left), and differenced pattern (right)*

However, in the case of the differenced schedule, the spectrums are evenly distributed over all frequencies, thus the NCP shows a relatively straight line within 95% confidence limits (Figure 5) such as white noise. Based on the result, it can be concluded that the occupancy pattern of experiment B also follows the random walk. Again, likewise Exp. A, the occupancy pattern has periodicity macroscopically, variations are irregular microscopically.

DISCUSSION AND CONCLUSION

The building performance simulation is quite limited when used for existing buildings due to occupants. Especially, the accurate information of occupancy pattern is very important for reliable simulation.

This paper explored the occupancy pattern from the viewpoint of a random walk which is a sequence of random variables. A series of experiments was conducted to obtain occupancy data in two laboratories. For the laboratories, the students entered and left freely during working hours. The statistical test, the normalized cumulative periodogram, was conducted to test whether the occupancy pattern is randomly changing or not.

The result shows that the differenced schedules (Equation 2) of two experiments were random sequences and followed white noise. It means that there is strong ‘random walk’ evidence for the occupancy pattern. On the other hand, it implies that the reproducibility (or prediction) of the occupancy pattern of existing buildings is quite low. In other words, transition probability from a space to a space of each individual is not constant because existing buildings are a part of the social-ecological system (Sowa 1994, Oreszczyn 2004, Lu et al 2010a, Cole et al 2010). Our result may explain the weakness of the Markovian framework for generalization of occupant models. The concept of the Markovian framework in occupant behavior in BPS may not become a remedy.

ACKNOWLEDGEMENTS

This research is supported by a grant (11 Urban-Architecture G02) from Urban Architecture Program funded by the Ministry of Land, Infrastructure and Transport of Korean government.

REFERENCES

Cole, R.J., Brown, Z. and McKay, S. 2010. Building human agency: a timely manifesto, *Building Research & Information*, Vol.38, No.3, pp.339-350

Gelb, A., 1974. *Applied optimal estimation*. MIT Press.

Hipel K.W. and McLeod, A.I. 1994. *Time series modelling of water resources and environmental systems*. Elsevier.

IBPSA. 1987-2013. *Proceedings of the IBPSA (International Building Performance Simulation Association) conference* ('87. '91, '93, '95, '97, '99, '01, '03, '05, '07, '09, '13).

Lu, X., Clements-Croome, D. and Viljanen, M. 2010a. Integration of chaos theory and mathematical models in building simulation Part I: Literature review. *Automation in Construction*, 19, 447-451.

Mahdavi, A., Mohammadi, A., Kabir, E. and Lambeva, L. 2008. Occupants' operation of lighting and shading systems in office buildings. *Journal of Building Performance Simulation*, 1 (1), 57-65.

Oreszczyn, T. 2004. Our innate ability to think of new ways to use energy. *Energy & Environment*, Vol. 15, No. 6, pp.1011-1014

Page, J., Robinson, D., Morel, N. and Scartezzini, J.L. 2008. A generalised stochastic model for the simulation of occupant presence. *Energy and Buildings*, 40 (2), 83-98.

Sowa, J. 1994. Man as a stochastic factor in indoor air quality predictions, *Indoor Environment*, Vol.3, pp.119-122