



OPEN

Categorizing SHR and WKY rats by chi2 algorithm and decision tree

Ping-Rui Tsai^{1,5✉}, Kun-Huang Chen^{2,5}, Tzay-Ming Hong^{1✉}, Fu-Nien Wang³ & Teng-Yi Huang⁴

Classifying mental disorder is a big issue in psychology in recent years. This article focuses on offering a relation between decision tree and encoding of fMRI that can simplify the analysis of different mental disorders and has a high ROC over 0.9. Here we encode fMRI information to the power-law distribution with integer elements by the graph theory in which the network is characterized by degrees that measure the number of effective links exceeding the threshold of Pearson correlation among voxels. When the degrees are ranked from low to high, the network equation can be fit by the power-law distribution. Here we use the mentally disordered SHR and WKY rats as samples and employ decision tree from chi2 algorithm to classify different states of mental disorder. This method not only provides the decision tree and encoding, but also enables the construction of a transformation matrix that is capable of connecting different mental disorders. Although the latter attempt is still in its fancy, it may have a contribution to unraveling the mystery of psychological processes.

What are the roots of mental disorder¹ In the path to answering this question, researchers are beginning to untangle the common biology that links supposedly distinct psychiatric conditions. Along this line of efforts, the purpose of this study is to determine whether the combination of power law and decision trees would improve the efficiency at classifying different mental disorders or states. In light of such a concern, this article comprises two steps: (a) providing a method to encode fMRI to the power-law distribution; (b) using decision trees to classify encoding to correct state. Recent studies showed that human brain activity can be expressed by different network equations² with the aid of graph method by fMRI samples³. Power law has been reported for many complex physical systems. Examples are the city population⁴, world wide web⁵, fluctuations in financial market⁶ et al. In this paper, fMRI information are encoded by power law distribution. we discuss the power law trait of authorized⁷ rat samples with SHR and WKY. Besides, Isoflurane (ISO) is used to further divide our samples into four states: high isoflurane WKY = HW, high isoflurane SHR = HS, low isoflurane SHR = LS, and low isoflurane WKY = LW. The format of our sample is 11 slices, 525 times, 64 × 64, FOV = 30 mm, and slice thickness = 1 mm. And each state contains 20 data. mental disorder is regarded as the mental difference, isoflurane is regarded as stimuli.

We use the same procedure as Ref.¹ to obtain the power-law network distribution for the fMRI data of rats. Pearson correlation defined in Eq. (1) plays an important role because this calculable quantity can reflect the strength of positive correlation between any two voxels:

$$r(x_1, x_2) = \frac{[\overline{v(x_1, t)v(x_2, t)} - \overline{v(x_1, t)} \times \overline{v(x_2, t)}]}{[\sigma(x_1)\sigma(x_2)]} \quad (1)$$

where a voxel at position x and time t is denoted as $v(x, t)$ while σ represents the standard deviation. When Eq. (1) exceeds a threshold value, chosen to be 0.7, these two voxels are regarded as being linked. After statistically analyzing all different degree of connectivity in the whole brain, we can obtain the power law distribution. Figure 1A shows the average distribution of these four states and Fig. 1B,C and F are average distribution of two states. Power law distribution is extremely useful because it sheds light on the difficult problem of analysing mental states. But it is insufficient to merely use its exponent to distinguish samples because ROC index is always less than 0.7. Based on this reason, chi2 algorithm will be quoted to help us select the significant difference from a group of degrees.

After chi2 algorithm. C4.5 decision trees can produce a tree structure. Chi2 algorithm is one of them of algorithm to make tree, ID3 and ID5 and so on is popular too, Fig. 1D,E,G and H are distribution made by chi2 algorithm. There are significant differences between groups. In order to test ROC of tree. We use tenfold validation to avoid over-fitting. Our purpose is to demonstrate whether C4.5 decision trees can offer a better way to

¹Department of Physics, National Tsing Hua University, Hsinchu 30013, Taiwan, ROC. ²Present address: College of Management and Design, Ming Chi University of Technology, New Taipei City 243303, Taiwan, ROC. ³Department of Biomedical Engineering and Environmental Sciences, National Tsing Hua University, Hsinchu 30013, Taiwan, ROC. ⁴Department of Electrical Engineering, National Taiwan University of Science and Technology, Taipei 10607, Taiwan, ROC. ⁵These authors contributed equally: Ping-Rui Tsai and Kun-Huang Chen. ✉email: ron104022509@gmail.com; ming@phys.nthu.edu.tw

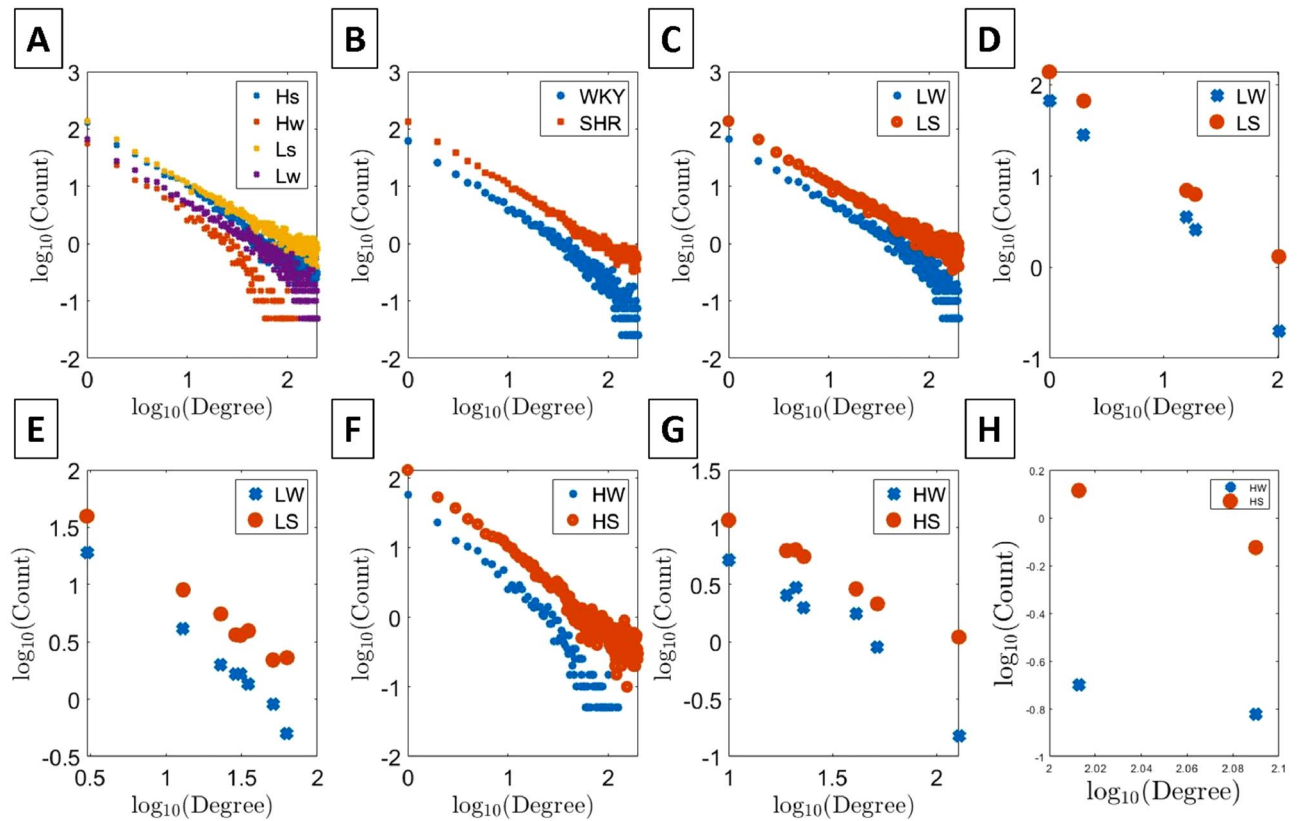


Figure 1. Power-law distribution consists of degree and its counted number. Each state, such as HS, HW, LS and LW, has 20 samples. Plots (A), (B), (C) and (F) are the distribution before chi2 algorithm. In (B), WKY and SHR are plotted without considering the anesthesia factor, and each distribution is deduced from 40 samples. In contrast, plots (D), (E), (G), and (H) are the distribution after chi2 algorithm.

help us observe and detect power law. Decision trees is a very popular tool for classification in data mining, which is widely used in deep-learning and machine learning^{8,9}, industrial application^{10,11}, medical treatments^{12–14}, and bioinformatics^{15–17}. To familiarize the readers with how decision trees can be of use in practical problems, let's imagine if we want to know who was dead among the passengers who boarded the Titanic. First, we can quote the list of passenger, such as gender, age or level of class on the boat. Second, using this list to make decision trees. Finally, decision trees will tell us which condition can effect the fate of each passenger. Obviously, in this paper power law is the analogy of the list - different degrees are like the factors. In the Result section of this paper, we will ensure the relation between C4.5 decision trees and the encoding fMRI of power law, such as a bar code and detector. This relation is rule how to readout sample condition, see Fig. 2A. In addition to this, this relation can help us to face a psychological processing which Encoding will be changed during dynamical, see Fig. 2B. in this project, we use MATLAB to deal with fMRI raw data. The GPU and CPU mixing program allows us to increase the execution efficiency. The former transforms the 4D (3D voxel space and time) into a 2D matrix (voxel position and time), while the latter handles the calculation of correlation. As for the decision trees, we save the MATLAB matrix file by csv format, input the data into Excel, and output to Java to build the decision trees for the final calculation of the outcome of tenfold cross-validation. We construct the transform matrix (TM) by using linear regression to find the relation between two mental disorder states. For example, to obtain the TM from LS to LW, we treat LW as a target state and a dependent variable which is defined by averaging power law distributions of LW. Our data set consists of 15 power law distributions from LS, defined as the original state. The weights of 200 linear regression equations are then utilized to construct TM, as shown in Fig. 4A. The process of encoding and decision tree is arranged in the “Materials and methods” section.

Results

All the branches of decision tree are important features from chi2 algorithm, see Fig. 3, while Fig. 1D–H show the distribution for degree after chi2 algorithm. By comparing with Fig. 1A–C, one can see that Fig. 1D–H manage to widen the separation between distributions for degree. Now we can input important features into C4.5 by tenfold cross-validation before selecting the trees whose ROC approaches 0.9. Figure 3A explains how the four states are distinguished. Initially, data are categorized into two groups, based on whether the dosage of isoflurane ($iso \geq 2.0$). Figure 3B and C show results from different situations of high ios and low ios groups. Our data are processed by ten fold cross-validation to achieve good results with ROC > 0.8. We double-confirm that power law has high sensitivity to represent the rat problem.

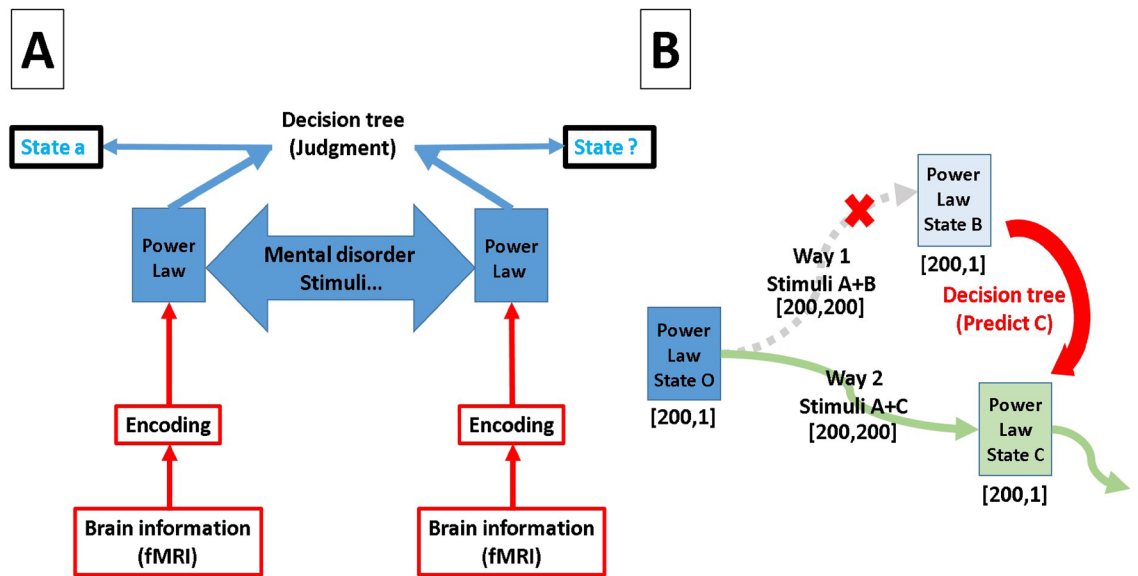


Figure 2. Plot (A) is a flow chart that includes encoding fMRI to power law and using decision tree to classify states. Plot (B) shows the transition between two states. If a situation has several paths, decision tree can help us detect which is the right way during evolution.

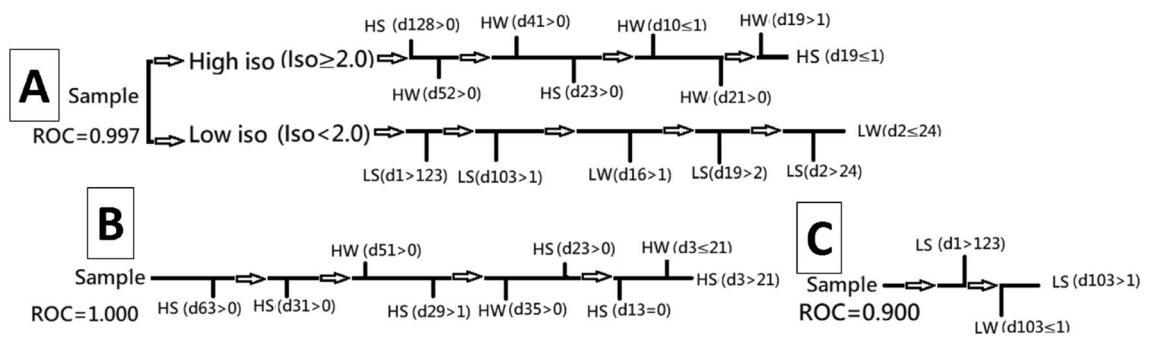


Figure 3. Plot (A) shows the decision tree of HS, HW, LS, and LW states. In contrast, plot (B) is of HS and HW states, and plot (C) of LS and LW states.

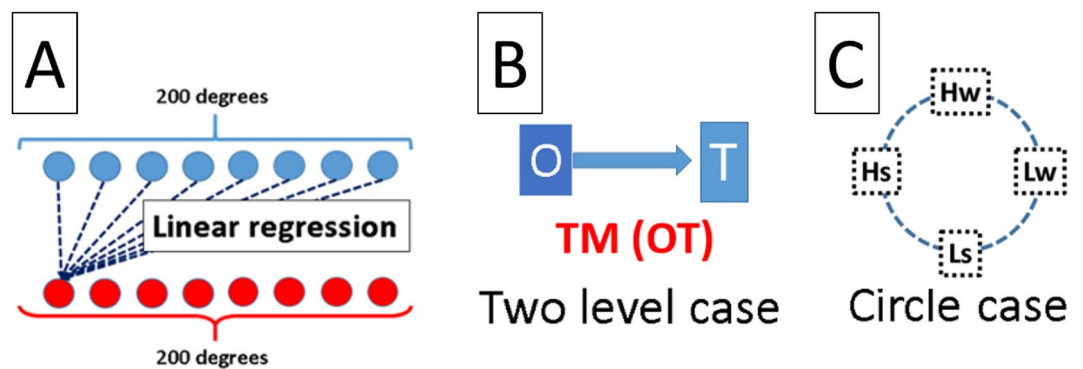


Figure 4. Plot (A) expresses TM constructed from linear regression between degrees. Plots (B) and (C) show the simulation about two level transition and four level transition.

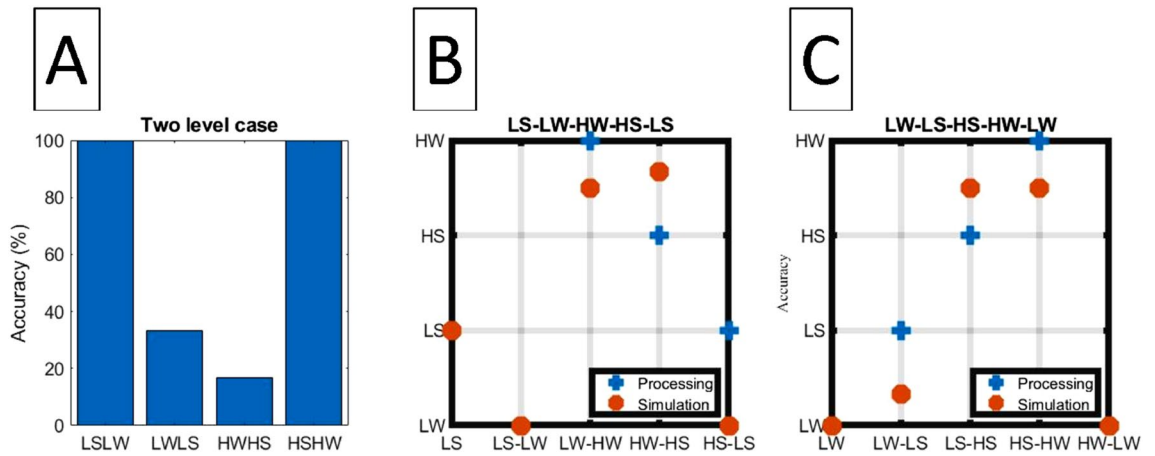


Figure 5. Plot (A) shows the accuracy in the case of two level transition from 5 testing samples. Compare with LSLW and HSHW that have a complete transition after detection by decision tree, the accuracy for LWLS and HWHS is lower. Several factors must be solved to obtain TM. Plots (B) and (C) show the circle case testing in Fig. 4B.

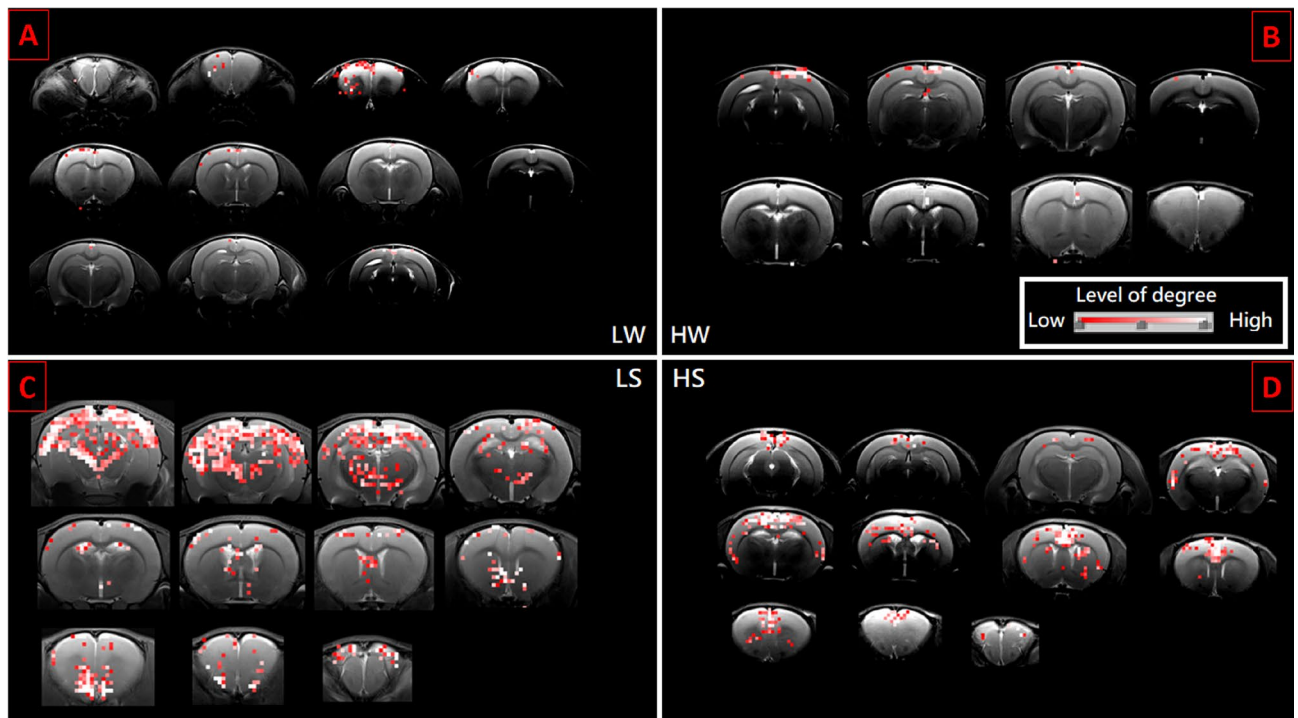


Figure 6. Panels (A) to (D) illustrate the activated region in LW, HW, LS, and HS states, respectively. The grade of brightness signifies different degrees of connectivity in the cross section. We only show the brain section with points.

Schematic plot in Fig. 4B describes the relation between original and target states which we regard as a two-level case. After the transformation via TM and being judged by decision tree, the accuracy is shown in Fig. 5A. Figure 4C describes that four TM are applied for each original state in order to obtain the results in Fig. 5B and C.

Discussion

Figure 6 shows that all states have different distributions for degree. Only the most representative result is selected among 20 rat samples for each state. In general, SHR/low-ISO has more degrees and covers more brain regions than WKY/ high-ISO. The rank 1 in HS is the Primary motor cortex, while HW, LS, and LW the Secondary motor cortex. In the mean time, the rank 2 in HS is Retrosplenial agranular cortex, while HW, LS, and LW are Primary motor cortex. Whenever the brain develops disorder, it exhibits a different functional network that consequently gives rise to a new exponent for the power law in Fig. 1. This is similar to the finding of Ref.² that the exponent may vary as the trial subject engages in different activities. We can find that the most active

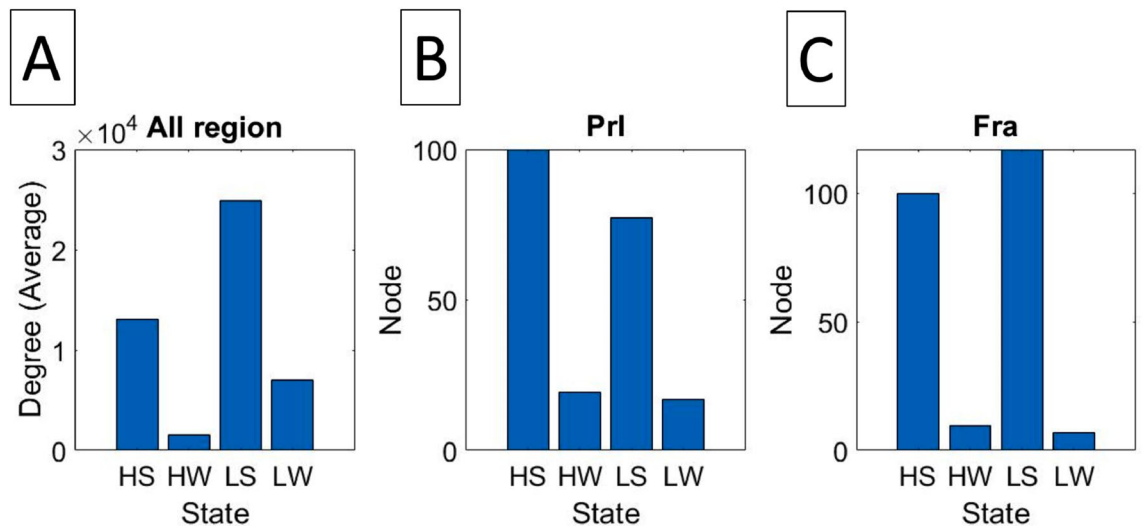


Figure 7. Plot (A) shows the total count of degree for the whole brain. Each state consists of 20 samples. (B) and (C) show Prl and Fra regions where the difference between SHR and WKY is significant. SHR has more nodes than WKY. Each state has 6 samples.

brain regions are the same. However, if we just focus on the samples of SHR, we discover that the secondary motor¹⁸ will fall to rank four. This is the reason why LS rat is more active than HS rat. More details will rely on more biological experiments in the future. The prefrontal cortex of ADHD patients has been reported to show abnormalities^{19,20}. In our case, we can check two important regions in the prefrontal cortex to distinguish different states. These regions, Prl^{21–24} and Fra²⁵, are related to the self-control and ADHD. The prefrontal cortex of SHR rat has been studied^{26–28}. Figure 7A shows the summation of degree over whole brain region with 6 samples in 4 states, while Fig. 7B and C show the average number of node in Prl and Fra. Note that the stimulus interaction for LS being lower than LW in Fra is contrary to our expectation. Future biological experiments are needed to clarify the source of this problem.

One common method in RS-fMRI (Resting state-fMRI) is Seed-based Correlation Analysis (SCA). SHR and WKY have been studied⁷. In general, SCA requires the choice of a Region Of Interest (ROI) to be a priori assumption, and needs to average over seed regions before calculating connectivity. In contrast, we forgo this step to avoid any subjective and abnormal seed from affecting the outcome. Instead, we upgrade and simplify the graph theory by numerical representations. Whether this approach is applicable to all states of mental disorder and has restrictions are important questions to answer in the future. In this work we have managed to establish that the power-law distribution carries enough information to deal with the trial subjects in this case. To locate the characteristics of any mental state, one only needs to use chi2 algorithm to pick out important degrees filtered from the power-law distribution.

In the past, researchers found that brains have two large imposing systems in the resting state. One is the DMN (Default mode network), while the other is composed of attentional or task-based systems²⁹. This motivates us to check whether double power laws may turn out to describe the distribution for degree better than the usual simple power law. In other words, can it be that each of these two systems contributes independently and gives rise to two different exponents? AIC is a statistical method for distinguishing the best fitting function among multiple candidates. Basically it balances the principles of accuracy (i.e., minimum loss of information) and frugality. Here we show the outcome of four states and human resting-state by AIC (Akaike information criterion)³⁰. Respectively, the AIC values calculated for single/double power law are (1) 729564/722132 for healthy humans in resting state, (2) 4016.756/4020.304 in HS, (3) 1036.135/1025.771 in HW, (4) 2265.964/2266.234 in LW, and (5) 5904.362/5908.362 in LS. Based on this information, we conclude that (1) healthy humans in resting-state for which double power laws fit better, (2) single power law wins out by a small margin for low iso rats. It is worth noting that this result should be treated with cautions because LW and LS rats are hard to remain still during fMRI scanning, (3) high iso WKY rats also favor the double powers, and (4) single power law is a better fit for SHR rats. Recent studies found that SHR (ADHD) children usually exhibit abnormal DMN network³¹. It has also been reported that mental disorder such as Alzheimer^{32,33}, depression³⁴, Schizophrenia³⁵, and ASD³⁶ can render DMN abnormal. It remains a pressing task to clarify whether the transition of double powers to single power correlates with abnormal DMN. In summary, power-law distribution can not only reflect the mental condition of our samples, but also reveal detail information about their network properties.

It is desirable to have more samples to maximize our use of Decision Tree to select power law from MRI data of mentally disordered rats. Although our results have demonstrated that the power-law distribution can be analyzed by Decision Tree to classify dosage of ISO and SHR vs. WKY, to vindicate its versatility more promotions are needed, e.g., depression, hypertension, or transient ischemic attack. We have two ideas to improve current understandings of the dynamics of brain: first, establish a relationship between observers (i.e., Decision Tree) and the objects being observed (i.e., power law from different states.). Once this relationship is available, it may function as a starting point to reveal possible connections among different observed objects.

In the TM, several problems remain to be solved: (1) different group samples may exhibit different ISO, age and so on. Although TM has enjoyed success in some transitions, (2) we yet need to determine whether TM is a linear or nonlinear system. Finally, (3) what are the relations between TM and stimuli is also an urgent question.

Materials and methods

Process. Step 1: Using Pearson correlation to calculate fMRI (spatial and temporal dimension) to power law distribution. Step 2: Using chi2 algorithm to select important features from degrees and making tree structure during tenfold. Afterwards, outputting the tree with the best ROC.

Animals: rat. All the experimental animals were admitted by the National Tsing Hua University Institutional Animal Care and Use Committee and complied with experimental guidelines. (https://drive.google.com/file/d/17cHSrbvBxaqpvEb_b-huZahQ7Ti3aqjL/view?usp=sharing).

Human subjects. We used resting-state fMRI images collected from six normal human subjects to test single and double power law. The datasets were downloaded from the ADHD-200 Consortium (http://fcon_1000.projects.nitrc.org/indi/adhd200). All MRI scans of the datasets were performed in the New York University Child Study Center³⁷. The scan protocol was approved by the institutional review boards of the New York University School of Medicine and New York University. The informed consents were obtained from all subjects. The protocols were performed in accordance with HIPAA guidelines and 1000 Functional Connectomes Project protocols. Our retrospective study using the ADHD-200 database was approved by the institutional review boards of National Taiwan University Hospital. The ADHD-200 Consortium provided de-identified datasets and removed the protected health information.

Magnetic resonance imaging. Our raw data come from the same source in Ref.²⁹, in this section, copyright is owing to the author who wrote “Magnetic resonance imaging” in the Ref.¹⁵. We scanned all animals with 7-Tesla Bruker Clinscan, which had a volume coil for signal excitation and a brain surface coil for signal receiving. The anesthesia process is operated by 1.4–1.5% isoflurane mixed with O₂ at flow rate of 1 L/min. We monitored all rats, made sure the respiratory rate in the range of 65–75 breaths/min while the scanning period, and body temperature maintained at 37 °C by a temperature-controlled water circulation machine. During the rs-fMRI experiments, we used gradient echo echo-planar-imaging (EPI) getting the 300 consecutive volumes with 11 coronal slices. The EPI specification is TE/TR = 20 ms/1000 ms, matrix size is 64 × 64, FOV = 30 × 30 mm² and slice thickness = 1 mm. We get the anatomical images by turbo-spin-echo (TSE) with scanning parameters of TE/TR = 14/4000, matrix size = 256 × 256, FOV = 30 × 30 mm², slice thickness = 1 mm, number of average = 2. To inspect the result of deep anesthesia, we applied 2.5–2.7% isoflurane mixed with O₂, and monitored respiratory rate in the range of 40–45 breaths/min during the whole scanning period.

Data processing for distribution for degree. Here we analyze our raw data from fMRI Grayscale image. Afterward, we transform them to scalar value matrix by MATLAB 2015a and 2018 version. This matrix is four dimensional, 64 × 64 × 11 × 525 where the first three components denote spatial position, while the last component refers to the time section in the scanning. Four dimensions render the matrix hard to manipulate, and it costs a lot of computer time. One can use GPU computing to disassemble it to two-dimensional form (45,056 × 525). After using Eq. (1) to calculate all voxels of degree in whole brain, we can get the distribution for degree for any sample. We calculate all 80 rats (each state has 20 samples) in order to obtain the form of the distribution for degree.

Chi2 algorithm. The feature selection on this study stems from chi2 algorithm³⁸ which is designed to discretize numeric attributes based on the X^2 statistic, and consists of two phases. In the first phase, it begins with a high significance level (sigLevel), Phase 1 is, as a matter of fact, a generalized version of ChiMerge of Kerber. Phase 2 is a ner process of Phase 1. Starting with sigLevel0 determined in Phase 1, each attribute i is associated with a sigLevel[i], and takes turns for merging. Consistency checking is conducted after each attribute's merging. At the end of Phase 2, if an attribute is merged to only one value, it simply means that this attribute is not relevant in representing the original data set. As a result, when discretization ends, feature selection is accomplished.

Data processing for C4.5 decision tree. In this study, we promote a set of new algorithms to enhance the Identifying effectiveness of SHR and WKY. The classifier algorithms are a combination of chi2 algorithm and C4.5 decision tree (C4.5), the chi2 algorithm evaluates the worth of a subset of attributes and C4.5 speculate the mental disorder. The chi2 algorithm is commonly used for testing relationships between categorical variables. The calculation of the chi2 algorithm is follows $X^2 = \sum \frac{(f_o - f_e)^2}{f_e}$, where f_o = the observed frequency (the observed counts in the cells) and f_e = he expected frequency if NO relationship existed between the variables. The decision tree algorithm is well known for its robustness and learning efficiency with a learning time complexity of $O(n \log 2n)$ ³⁹. C4.5 has been listed in the item 10 algorithms in data mining⁴⁰. It is a popular statistical classifier developed by Ross Quinlan in 1993. Basically, C4.5 is an extension of Quinlan's earlier ID3 algorithm. In C4.5 the Information Gain split criterion is replaced by an Information Gain Ratio criterion which penalizes variables with many states. C4.5 can be used to generate a decision tree for classification. The learning algorithm applies a divide-and-conquer strategy⁴¹, to construct the tree. The sets of instances are accompanied by a set of genes (attributes). This classifier has additional features, such as handling missing values, categorizing continu-

ous attributes, pruning decision trees, deriving rules, endotestae Information gain (S, A) of a feature A relative to a collection of examples S, is defined as $Gain(S, A) = Entropy(S) - (\sum_{S_V} \frac{|S_V|}{|S|} \times Entropy(S_V))$, where Values (A) is the set of all possible values for attribute A, and S_V is the subset of S for which feature A has value v (i.e. $S_V = \{s \in S \mid A(s) = v\}$), Note the first term in the equation for Gain is just the entropy of the original collection S and the second term is the expected value of the entropy after S is partitioned using feature A. The expected entropy described by the second term is the direct sum of the entropy of each subset S_V , weighed by the fraction of samples $\frac{|S_V|}{|S|}$ that belong to S_V , Gain (S, A) is therefore the expected reduction in entropy caused by knowing the value of feature A. The Entropy is given by $Entropy(S) = \sum_{i=1}^n -P_i \log(P_i)$.

Data processing for testing single and double power law. We choose the same method and procedures described in Ref.⁴¹. Details can be found in Sec.IV⁴².

Data processing for calculating L. When we calculated the degree for any voxels, the path length (L) between two voxels is defined as the minimum number of links necessary to connect each other. Also, we collect L from data of six rats for each state at the same time. Afterward, MATLAB 2018a is employed to find the max path and average L.

Data processing for C and structure network patterns. First, we print out all connection information between two voxels for any sample in txt format. Then, insert the data in MATLAB 2018a to obtain the functional brain network pattern. If interested at calculating the clustering coefficient for any voxel linked, you have to get information of degree and the number of links connecting the neighbors. Finally, the average C can be determined from this equation $C = \frac{1}{N \sum_{i=1}^N C_i}$ where N is the number of voxels and i the voxel number.

Received: 8 January 2020; Accepted: 18 January 2021

Published online: 10 February 2021

References

1. Marshall, M. The hidden links between mental disorders. *Nature* **581**, 19–21 (2020).
2. Egulluz, V. M. *et al.* Scale-free brain functional network. *Phys. Rev. Lett.* **94**, 018102 (2005).
3. Maslov, S., Sneppen, K. & Alon, U. In *Handbook of Graphs and Networks* (eds Bornholdt, S. & Schuster, H. G.) (Wiley, Weinheim, New York, 2003).
4. Gabaix, X. Zipf's law for cities: an explanation. *Q. J. Econ.* **114**, 739–767 (1999).
5. Albert, R. *et al.* Internet: diameter of the world-wide web. *Nature* **401**, 130–131 (1999).
6. Gabaix, X. *et al.* A theory of power-law distributions in financial market fluctuations. *Nature* **423**, 267–270 (2003).
7. Huang, S.-M. *et al.* Inter-strain differences in default mode network: a resting state fMRI study on spontaneously hypertensive rat and Wistar Kyoto Rat. *Sci. Rep.* **6**, 21697 (2016).
8. Zen, H. *et al.* Statistical parametric speech synthesis using deep neural networks. In *IEEE International Conference on Acoustics, Speech and Signal Processing* 7962–7966 (2013).
9. Fayyad, U. M. *et al.* On the handling of continuous-valued attributes in decision tree generation. *Mach. Learn.* **8**, 87–102 (1992).
10. Sheng, Y. *et al.* Decision tree-based methodology for high impedance fault detection. *IEEE Trans. Power Deliv.* **19**, 533–536 (2004).
11. Kumar, R. *et al.* Recognition of power-quality disturbances using S-transform-based ANN classifier and rule-based decision tree. *IEEE Trans. Ind. Appl.* **51**, 1249–1258 (2015).
12. Shouman, M. *et al.* Using decision tree for diagnosing heart disease patients. In *Proceedings of the 9th Australasian Data Mining Conference (AusDM'11)*, Vol. 121, 23–29 (2011).
13. Pavlopoulos, S. A. *et al.* A decision tree-based method for the differential diagnosis of aortic stenosis from mitral regurgitation using heart sounds. *Biomed. Eng. Online* **3**, 21 (2004).
14. Qu, Y. *et al.* Boosted decision tree analysis of surface-enhanced laser desorption/ionization mass spectral serum profiles discriminates prostate cancer from noncancer patients. *Clin. Chem.* **48**, 10 (2002).
15. Chen, K.-H. *et al.* Gene selection for cancer identification: a decision tree model empowered by particle swarm optimization algorithm. *BMC Bioinform.* **15**, 15–49 (2014).
16. Dima, S. *et al.* Decision tree approach to the impact of parents' oral health on dental caries experience in children: a cross-sectional study. *Int. J. Environ. Res. Public Health* **15**, 692 (2018).
17. Chen, K.-H. *et al.* Applying particle swarm optimization-based decision tree classifier for cancer classification on gene expression data. *Appl. Soft Comput.* **24**, 773–780 (2014).
18. Sul, J. H. *et al.* Role of rodent secondary motor cortex in value-based action selection. *Nat. Neurosci.* **14**, 1202–1208 (2011).
19. Molupe, M. *et al.* NMDA receptor function in the prefrontal cortex of a rat model for attention-deficit hyperactivity disorder. *Metab. Br. Dis.* **19**, 35–42 (2004).
20. Khorevin, V. I. *et al.* Effect of L-DopA on the behavioral activity of wistar and spontaneously hypertensive (SHR) rats in the open-field test. *Neurophysiology* **36**, 116–125 (2004).
21. dela Peña, I. J. I. *et al.* Transcriptional profiling of SHR/NCrl prefrontal cortex shows hyperactivity-associated genes responsive to amphetamine challenge. *Genes Br. Behav.* **16**, 664–674 (2017).
22. Bymaster, F. P. *et al.* Atomoxetine increases extracellular levels of norepinephrine and dopamine in prefrontal cortex of rat: a potential mechanism for efficacy in attention deficit/hyperactivity disorder. *Neuropsychopharmacology* **27**, 699–711 (2002).
23. Russell, V. A. *et al.* Hypodopaminergic and hypernoradrenergic activity in prefrontal cortex slices of an animal model for attention-deficit hyperactivity disorder—the spontaneously hypertensive rat. *Behav. Br. Res.* **130**, 191–196 (2002).
24. Duan, C. A. *et al.* Requirement of prefrontal and midbrain regions for rapid executive control of behavior in the rat. *Neuron* **86**, 1491–1503 (2015).
25. Hambrecht, V. S. *et al.* G proteins in rat prefrontal cortex (PFC) are differentially activated as a function of oxygen status and PFC region. *J. Chem. Neuroanatomy* **37**, 112–117 (2009).
26. Russell, V. *et al.* Altered dopaminergic function in the prefrontal cortex, nucleus accumbens and caudate-putamen of an animal model of attention-deficit hyperactivity disorder—the spontaneously hypertensive rat. *Br. Res.* **676**, 235–424 (1995).
27. Russell, V. *et al.* Increased noradrenergic activity in prefrontal cortex slices of an animal model for attention-deficit hyperactivity disorder—the spontaneously hypertensive rat. *Behav. Br. Res.* **117**, 69–74 (2000).
28. Russell, V. A. Increased AMPA receptor function in slices containing the prefrontal cortex of spontaneously hypertensive rats. *Metab. Br. Dis.* **16**, 143–149 (2001).

29. Bullmore, E. & Sporns, O. Complex brain networks: graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* **10**, 186–198 (2009).
30. Akaike, H. In Proceedings of the Second International Symposium on Information Theory (ed. Petrov, B. N. and Csaki, E.) (Akademiai Kiado, Budapest, 1973), pp. 267–281; *IEEE Trans. Autom. Control* **19**, 716 (1974); *A Celebration of Statistics* (ed. Atkinson, A. C. and Fienberg, S. E.) 1–24 (Springer, Berlin, 1985).
31. Liddle, E. B. *et al.* Task-related default mode network modulation and inhibitory control in ADHD: effects of motivation and methylphenidate. *J. Child Psychol. Psychiatry* **52**, 761–771 (2011).
32. Mevel, Katell *et al.* The default mode network in healthy aging and Alzheimer's disease. *Int. J. Alzheimer's Dis.* 535816 (2011).
33. Jones, D. T. *et al.* Age-related changes in the default mode network are more advanced in Alzheimer disease. *Neurology* **77**, 1524–31 (2011).
34. Shelton, Y. I. *et al.* The default mode network and self-referential processes in depression. *PNAS* **106**, 1942–1947 (2009).
35. Salgado-Pineda, P. *et al.* Correlated structural and functional brain abnormalities in the default mode network in schizophrenia patients. *Schizophr. Res.* **125**, 101–109 (2011).
36. Starck, T. *et al.* Resting state fMRI reveals a default mode dissociation between retrosplenial and medial prefrontal subnetworks in ASD despite motion scrubbing. *Front. Hum. Neurosci.* **7**, 802 (2013).
37. The ADHD-200 Consortium, *Front Syst Neurosci.* 2012; 6:62
38. Liu, H. & Setiono, R. *Chi2: feature selection and discretization of numeric attributes*. In Proceedings of 7th IEEE International Conference on Tools with Artificial Intelligence, 388–391 (Herndon, VA, 1995).
39. Tan, A. C. & Gilbert, D. Ensemble machine learning on gene expression data for cancer classification. *Appl. Bioinform.* **2**, 75–83 (2003).
40. Wu, X. *et al.* Top 10 algorithms in data mining. *Knowl. Inform. Syst.* **16**, 1–37 (2008).
41. Quinlan, J. R. *C4.5: Programs for Machine Learning* (Morgan Kaufmann, Los Altos, 1993).
42. Tsai, S.-T. *et al.* Power-law ansatz in complex systems: excessive loss of information. *Phys. Rev. E* **92**, 062925 (2015).

Acknowledgements

We acknowledge funding from the Ministry of Science and Technology in Taiwan under Grant No. 105-2112-M-007-008-MY3 and 108-2112-M007-011-MY3. We also thank Prof. Chung-Chuan Lo for helpful advices, Prof. Che-Rung Lee and Quey-Liang Kao for the access to their GPU workstation, and Kun-I Chao and Li-Jie Chen for helping us process images of fMRI.

Author contributions

Hosting Plan: P.-J.T. and K.-H.C. Coding program: P.-J.T. Image processing: P.-J.T. Data processing: P.-J.T. Section about Magnetic resonance imaging: author of Ref.²⁹. Statistical Analysis: K.-H.C. (for chi2 algorithm, C4.5 decision tree, t-test, and ROC), P.-J.T. (for AIC and power law). Academic guidance: F.-N.W., T.-M.H. and T.-Y.H. Writing: P.-J.T. and T.-M.H. (Abstract, Introduction, Result, Discussion), P.-J.T., K.-H.C., and author of Ref.²⁹ (“Materials and methods”).

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to P.-R.T. or T.-M.H.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2021