Mathematical modeling of the malignancy of cancer using graph evolution

Cigdem Gunduz-Demir *

Department of Computer Engineering, Bilkent University, Ankara TR-06800, Turkey

Abstract

We report a novel computational method based on graph evolution process to model the malignancy of brain cancer called glioma. In this work, we analyze the phases that a graph passes through during its evolution and demonstrate strong relation between the malignancy of cancer and the phase of its graph. From the photomicrographs of tissues, which are diagnosed as normal, low-grade cancerous and high-grade cancerous, we construct cell-graphs based on the locations of cells; we probabilistically generate an edge between every pair of cells depending on the Euclidean distance between them. For a cell-graph, we extract connectivity information including the properties of its connected components in order to analyze the phase of the cell-graph. Working with brain tissue samples surgically removed from 12 patients, we demonstrate that cell-graphs generated for different tissue types evolve differently and that they exhibit different phase properties, which distinguish a tissue type from another.
© 2007 Elsevier Inc. All rights reserved.

Keywords: Mathematical modeling of cancer malignancy; Mathematical cancer grading; Histopathological image analysis; Graph evolution; Graph theory; Glioma

* Tel.: +90 312 290 3443; fax: +90 312 266 4047.
  E-mail address: gunduz@cs.bilkent.edu.tr

1. Introduction

Real-world graphs [1] have been extensively studied to model different phenomena in man-made [2–4], social [5–9], and biological systems [10–13]. The graphs of man-made systems include the Internet router graph with routers being the nodes and fiber connections being the edges [2], the World Wide Web graph with web pages being the nodes and URLs being the edges [3], and the USA power grid network with generators, transformers, and substations being the nodes and high-voltage transmission lines being the edges [1]. The graphs of social systems include the Hollywood movie star network with actors being the nodes and co-starring in the same movie being the edges [1], scientific collaboration network with scientists being the nodes and co-authoring in the same publication being the edges [5], and the SARS disease spreading network with people being the nodes and contacts between any two people being the edges [6]. The graphs of biological systems include the neural network of the worm *Caenorhabditis elegans* with neurons being the nodes and neural connections being the edges [1], the protein–protein interaction graph of the bacteria *Saccharomyces cerevisiae* with proteins being the nodes and direct physical interactions being the edges [10], and metabolic graphs of different organisms with substrates being the nodes and actual metabolic reactions being the edges [11]. Obviously, all these systems are self-organizing structures and have very different domain and sizes. However, these systems exhibit common classes of topological properties which are quantified with the definition of graphs.

In this paper, we propose to use graphs to model the malignancy of brain cancer called glioma. From the photomicrograph of a tissue, we construct a cell-graph by defining nodes as individual cells within the tissue and generating edges between these nodes depending on the Euclidean distance between them. The proposed method relies on modeling the changes in the cell distribution of a tissue that are caused by cancer as well as modeling the degree of these changes. For that, it employs graph evolution process, which is quantified with the definition of graph phases. In this work, we gradually increase the connectivity level of a graph such that it evolves from an empty graph to a connected graph (i.e., from having isolated nodes to having all nodes being connected to each other) and analyze the phases that the graph passes through during its evolution. Working with brain tissues of 12 different patients, which are diagnosed as normal, low-grade cancerous (gliomatous), and high-grade cancerous (gliomatous), we demonstrate that cell-graphs generated for different types of tissues evolve differently. Moreover, we show that this difference can be used to differentiate the tissue types, i.e., to distinguish normal, low-grade gliomatous, and high-grade gliomatous tissues from each other. To the best of our knowledge, this is the first demonstration of the use of graph evolution process for the purpose of modeling the malignancy of cancer. This demonstration opens up possibilities to quantitatively model the dynamics of cancer and simulate its progress using graph evolution.

In our previous studies [14–16], we propose to use cell-graphs for the purpose of cancer diagnosis. In those studies, we construct simple [14,15] and weighted [16] graphs from tissues and compute their local and global graph metrics in order to distinguish cancerous tissues (regardless of their malignancy grades) from non-cancerous ones. As opposed to our previous studies, this work proposes a novel computational model based on the use of graph evolution process for the purpose of cancer grading. The correct grading of cancer is very important since the selection of treatment largely relies on this grading as well as the malignancy grade is an important predictor for the survival time of a patient. Nevertheless, there may exist a considerable amount of intra- and

inter-observer variability in cancer grading [17,18]. Therefore, mathematical modeling provides an objective measure to help to reduce this variability.

The remaining of the paper is organized as follows: In Section 2, we briefly describe the methodology to generate a cell-graph from a tissue image as well as the details of the connectivity analysis to determine the phase of the generated cell-graph. In Section 3, we present our experimental results. We discuss our results and provide a summary of our work and a future perspective for our research in Section 4.

2. Methodology

2.1. Cell-graph generation

We construct a cell-graph from a tissue image by segmenting the individual cells as nodes and probabilistically assigning edges between these nodes depending on the Euclidean distance between the cells.

In cell segmentation, we use a semi-automated method that consists of automated segmentation followed by manual segmentation. In the automated part, the first step is to distinguish the pixels of cell nuclei from those of the background. For that, we threshold the pixels of an image; we use the Otsu method [19] to determine an optimal threshold value. After thresholding, we eliminate small objects and break narrow strips making use of morphological operators [20]. At the end of this step, each pixel is identified as a "cell" or "background" pixel. For an exemplary tissue image (Fig. 1a), Fig. 1b illustrates "cell" and "background" pixels with black and white, respectively. In the second step of the automated part, we segment the individual cells by applying the watershed algorithm [21]. This algorithm may result in representing an individual cell with multiple points or representing multiple cells with a single point. If such a case occurs, we manually merge or split such points to obtain the final locations of individual cells. At the end of the manual segmentation, the centroids of the final locations of cells are considered to be the coordinates of nodes. Fig. 1c shows the centroids with green[1] dots.

After identifying the coordinates of nodes, the next step is to establish edges between these nodes in order to generate a cell-graph. The cell-graph approach relies on extracting information about the pairwise cell distances. However, instead of representing all of these distances, it represents only a group of them by probabilistically deciding the existence of each edge based on the Euclidean distance between its end points. Using only such representative distances would be especially important when large number of nodes is used.[2] In this work, the probability of assigning an edge between nodes u and v is defined as

$$P(u, v) = d(u, v)^\alpha \tag{1}$$

where $d(u,v)$ is the Euclidean distance between the nodes u and v and the edge exponent α is a model parameter that should be selected smaller than zero. In this work, we prefer this prob-

---

[1] For interpretation of the references to color in this figure, the reader is referred to the web version of this article.
[2] Our experiments show that promising results are obtained by representing <200 of such distances. This is much less than the number of all pairwise distances which is $O(N^2)$; in our experiments this number is >11000.
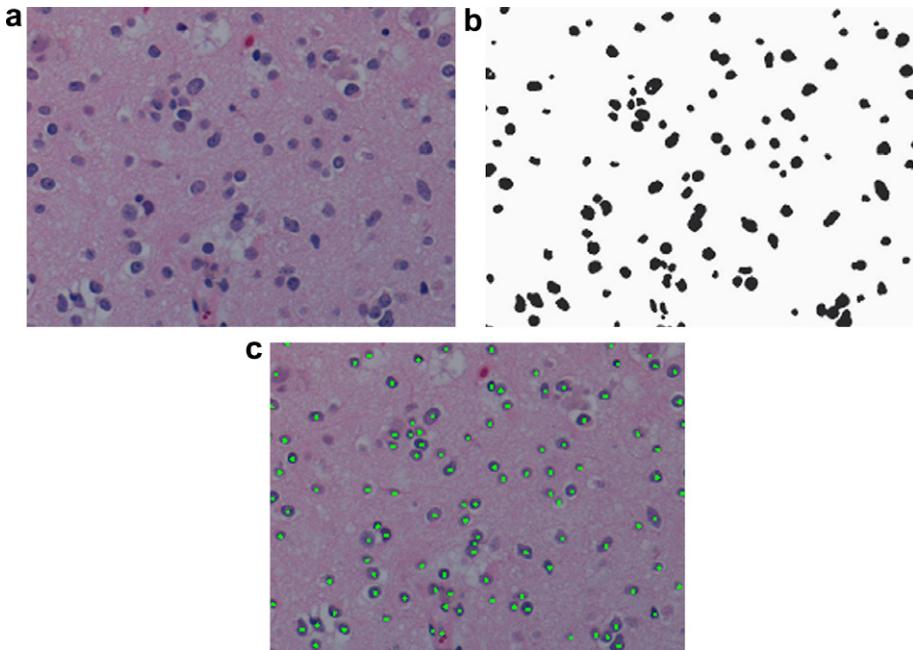
Fig. 1. Steps for the semi-automated cell segmentation: (a) start with the original tissue image, (b) threshold the pixels of the image and apply morphological operators, and (c) apply the watershed algorithm and if necessary, manually segment the cells.

ability function, which is also used in one of our earlier work [15], since it has a single control parameter so that we could control the graph connectivity by only this single parameter. In this function, the probability of being an edge between two nodes decays with the increasing Euclidean distance between these two nodes. In our earliest work [14], we also use another probability function (the Waxman model) that also assigns edges depending on the Euclidean distances between the nodes. However, the Waxman model has two control parameters. Moreover, it normalizes each Euclidean distance with the largest possible Euclidean distance between any two nodes. In diagnosis, this normalization yields equally accurate results since the cell distributions of malignant (regardless of the malignancy grade) and benign tissues are completely different. In the grading, such normalization may yield similar results for some of the cancerous tissues with different malignancy grades.

2.2. Connectivity analysis

Our technique relies on the use of graph evolution process to model the malignancy of cancer. For a tissue, we generate cell-graphs by gradually increasing the graph connectivity and quantify the characteristics of their evolution by the use of graph phases. In cell-graphs, the connectivity is determined by two factors: the spatial distribution of nodes (cells) and the edge exponent. As the cell distribution, which is specific to the tissue, cannot be controlled externally, we use the edge exponent to control the graph connectivity.

Table 1
Quantitative measures used in the phase analysis of a cell-graph

1. Number of edges
2. Number of connected components
3. Number of isolated nodes
4. Number of trees
5. Number of nodes belonging to trees
6. Number of components with equal number of nodes and edges
7. Number of nodes belonging to components with equal number of nodes and edges
8. Number of components with complex structures
9. Number of nodes belonging to components with complex structures
10. Size of the giant connected component
11. Structural type of the giant connected component

As the edge exponent increases, the generated cell-graphs evolve and pass through different phases. These phases correspond to different ranges of connectivity (i.e., number of edges) and they exhibit different characteristics. For cell-graphs, we define the following phases with different connectivity levels and characteristics, with N and E being the number of nodes and edges, respectively:

Phase 1: This phase corresponds to $E < \frac{N}{2}$. The components of the generated cell-graphs are almost surely either isolated nodes or trees or those consisting of equal number of nodes and edges (containing a cycle).

Phase 2: This phase corresponds to $\frac{N}{2} \leqslant E \leqslant \frac{N \log N}{2}$. When the number of edges exceeds the $N/2$ threshold, the structure of cell-graphs changes abruptly such that the giant connected components of these graphs have rather complex structures. The other graph components are relatively small and tend to melt into the giant component as the edge exponent (i.e., number of edges) increases. In this phase, the size of the giant connected component jumps from $\sim N^{2/3}$ (when $E \sim N/2$) to $N/2$ (when the threshold is exceeded).

Phase 3: This phase corresponds to $E > \frac{N \log N}{2}$. In this phase, the cell-graphs almost surely become connected.

In the definition of these phases, we make use of the edge thresholds for which the graphs start holding certain properties. As studied by Erdos and Renyi, the structure of random graphs changes as the number of edges increases [22–24]. Erdos and Renyi show that there exist some thresholds for graphs to hold certain properties. Among these properties are those that the structure of the graph suddenly changes and that the graph becomes connected. They prove that these two properties are hold when the number of edges is equal to $N/2$ and $(N\log N)/2$, respectively. Moreover, they show that the size of the giant connected component becomes $\sim N^{2/3}$ when the first property is hold. In this work, we make use of these thresholds to define the phases of cell-graphs for the purpose of modeling the cancer malignancy. These phases are similar for both random graphs and cell-graphs. However, as opposed to cell-graphs, the spatial distribution of nodes does not affect the connectivity of random graphs; with the same number of nodes, the edge probability is the only factor that determines the graph connectivity. Thus, random graphs generated using
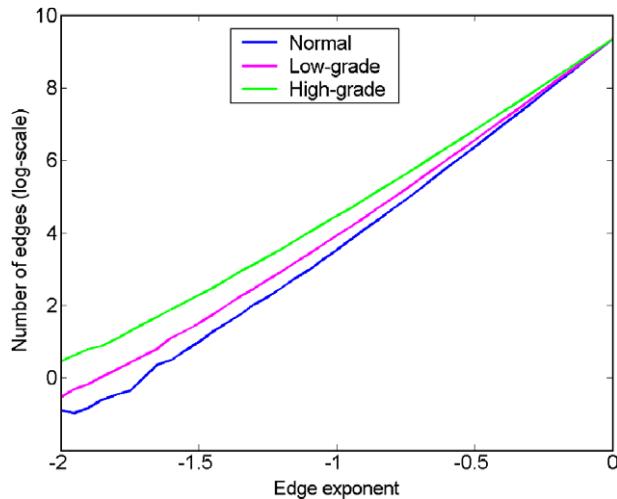
Fig. 2. Average number of edges for different tissue types. y-axis is plotted at log-scale to show the difference among different tissue types better.

the same number of nodes and the same edge probability are found in the same phase and, hence, random graphs with equal number of nodes pass through the same phases during their evolutions. On the other hand, with the same number of nodes and the same edge exponent, cell-graphs may lead to different phases depending on the spatial distribution of cells, which is specific to the tissue type. Therefore, cell-graphs of different tissue types evolve differently. We will thoroughly analyze the phases of cell-graphs generated for different tissue types (normal, low-grade cancerous, and high-grade cancerous) in Section 3.

3. Experiments

The experiments are conducted on photomicrographs of human brain tissue samples. Each sample consists of a 5–6 micron-thick tissue section stained with the hematoxylin–eosin technique and mounted on a glass slide. Images of these samples are taken in the RGB color space using a Nikon Coolscope Digital Microscope. Images are taken using a 20× microscope objective lens. In this experiment, we use 24 tissue samples taken from 12 different patients. These samples include both normal and cancerous (gliomatous) tissues. Cancerous tissues are differentiated as low-grade and high-grade. This sample set consists of 8 normal tissues from 4 patients, 8 low-grade cancerous tissues from 4 patients and 8 high-grade cancerous tissues from the remaining 4 patients.

For every tissue, we use an image containing the same number of cells in order to analyze the evolutions of their cell-graphs with the same number of nodes and the same edge exponent.[3] In our experiments, images consist of ∼150 cells, since the minimum number of cells in an image

---

[3] With the same number of nodes, random graphs pass through the same evolution process. On the other hand, depending on the spatial distributions of their nodes, cell-graphs evolve differently even with the same number of nodes. In order to examine this dependency, we fix the number of nodes for all tissue types.
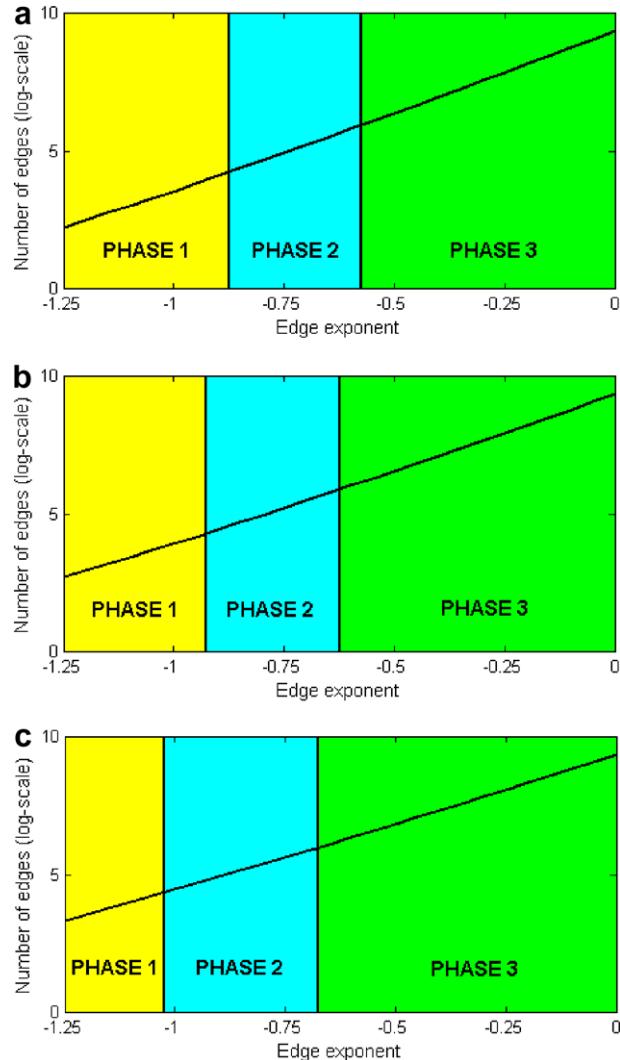
Fig. 3. Number of edges as a function of the edge exponent, for (a) normal tissues, (b) low-grade gliomatous tissues, and (c) high-grade gliomatous tissues. *y*-axes are plotted at log-scale to show the difference among different tissue types better.

is 152. On these images, we construct cell-graphs by identifying cells as nodes and assigning edges between these nodes depending on the probability distribution given in Eq. (1). In our experiments, we analyze the phase of cell-graphs by gradually increasing the connectivity. For that, we use an edge exponent, ranging from -2 to 0 ($\alpha = \{-2.00, \ -1.95, -1.90, \ldots, -0.10, -0.05, 0.00\}$). Using each edge exponent, we generate 10 different cell-graphs for each tissue image.

The phase of a cell-graph depends on the number of its edges E, the number of its connected components, and the properties of these connected components. In addition to using the distribution of the properties of all components, we separately use the properties of the giant con-
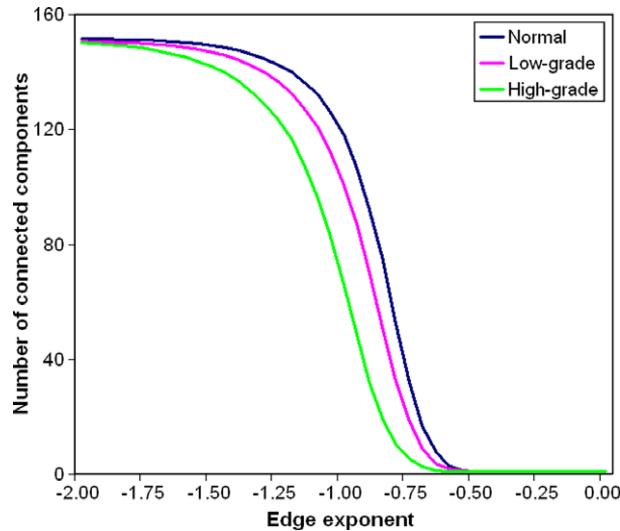
Fig. 4. Average number of connected components for different tissue types.

nected component. Among these properties are the number of nodes a connected component contains and its structural type. Depending on its structure, a connected component can be (i) an isolated node, (ii) a tree, (iii) a component with equal number of nodes and edges (containing a cycle), and (iv) a component with a complex structure. Note that a component has a complex structure if it is neither an isolated node, nor a tree, nor a component with equal number of nodes and edges. Table 1 summarizes these quantitative measures, which we use in the phase analysis of a cell-graph.

In Fig. 2, we report the average number of edges as a function of the edge exponent, for each tissue type[4]; to show the difference better among different tissue types, *y*-axis is plotted at log-scale. This figure shows that cell-graphs generated for different tissue types have different number of edges (and thus, different connectivity), although they are generated using the same number of nodes and the same edge exponent. This reveals that the connectivity of a cell-graph is determined by both the edge exponent and the spatial distribution of its nodes. Note that the connectivity of random graphs is only determined by the edge probability, i.e., almost the same amount of edges is generated with the same number of nodes and the same edge probability. In order to investigate whether or not the difference shown in Fig. 2 is significant, we use a *t*-test with a significance level of 0.05 for each of the given edge exponents. The *t*-test demonstrates that there is a statistically significant difference among the number of edges of the cell-graphs for different tissue types. The only exception occurs when the edge exponent is selected as 0.00; in this case, the number of edges for all tissues is the same since the probability of assigning an edge between any pair of nodes is equal to 1 and the number of nodes in all graphs is the same.

---

[4] In this work, for each tissue type, we compute the average over 80 different graph instances; there are eight images for each tissue type and we generate 10 different graphs for each tissue.

The phase of a graph is determined by the number of its edges together with the characteristics of its connected components. With approximately 150 cells (nodes) in a tissue (cell-graph), the edge thresholds are approximately 75 ($N/2$) and 375 ($N\log N/2$) for the phases that we explain in Section 2.2. In Fig. 3, we indicate the boundaries for which the phase transitions occur (i.e., the edge thresholds are exceeded); to determine these boundaries, we use the aforementioned phases (edge thresholds). This figure demonstrates that each tissue type has its own phase ranges, and thus, its own characteristic evolution process. As a result, it is possible to distinguish tissues of different types by making use of the phase information (evolution) of their cell-graphs.

Next, we analyze the connected components of cell-graphs. For each tissue type, Fig. 4 shows the average number of connected components as a function of the edge exponent. In this figure, the curve characteristics for all tissue types are consistent with the phase definitions. For edge exponents that correspond to Phase 1, the number of connected components is higher. In this range, first, the nodes are mostly disconnected since the edge exponent is too small to establish an edge between any pair of nodes. Therefore, the number of connected components is almost equal to the number of nodes. Then, with the increasing edge exponent, nodes are being connected, first forming trees of sizes 2–3 and then forming trees of higher sizes. In this range, the giant connected components consist of 15.19 nodes (normal), 20.71 nodes (low-grade glioma), and 21.26 nodes (high-grade glioma), on the average.

As the edge exponent increases (Phase 2), the probability of establishing an edge between a pair of nodes increases, leading to more number of edges established between both close and distant nodes. This results in larger number of nodes being connected in a single component. In this phase, for the edge exponent that leads to approximately 75 ($\sim N/2$) edges, the giant connected components consist of approximately 29 ($\sim N^{2/3}$) nodes. When this edge exponent is exceeded, the giant component size first goes to approximately 75 ($\sim N/2$) and then exceeds 100. Towards
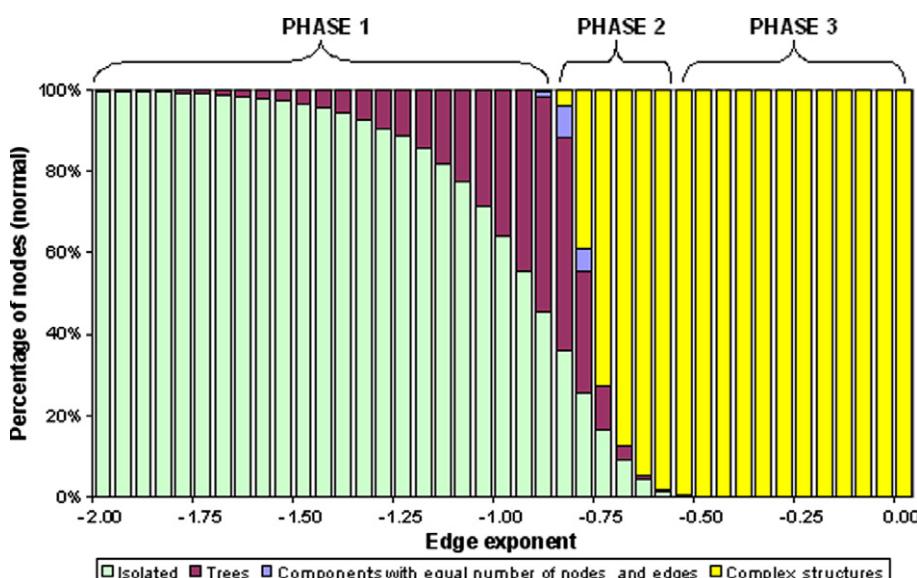
Fig. 5. For normal tissues, the distribution of nodes depending on the structure of the components that they belong to.

the end of this phase, the giant component size approaches to 150 ($\sim N$). This phase leads to giant connected components with more complex structures compared to trees. Besides these complex-structured components, others are mostly isolated nodes or trees with smaller sizes. These small components (isolated nodes and trees) tend to melt into the complex-structured giant component towards the end of this phase.

As the edge exponent further increases (Phase 3), the number of small components becomes smaller and smaller. In our experiments, it remains at most four isolated nodes at the beginning of this phase. As the edge exponent approaches to zero, all nodes are being connected to each other, and thus, the number of connected components becomes 1. In this range of the edge exponent, all giant connected components have more complex structures compared to trees. These observations are also illustrated in Figs. 5–7 for normal, low-grade cancerous (gliomatous) and high-grade cancerous (gliomatous) tissues, respectively. These figures show the phase ranges as well as the distribution of nodes depending on the structure of the components that they belong to.

For the results given in Figs. 4–7, we use a t-test with a significance level of 0.05 to investigate whether or not the difference is statistically significant. The t-test shows that for edge exponents smaller than -0.50, the difference among the number of connected components of different tissue types is statistically significant. For each exponent, we also compare the distribution of connected components in terms of their structural types. For that, among different tissue types, we compare the percentages of nodes that belong to each structural type. After performing t-tests for each structural type, we combine their results such that the difference in distributions is not statistically significant if there is not any difference in any of these t-tests. Likewise, our experiments show that
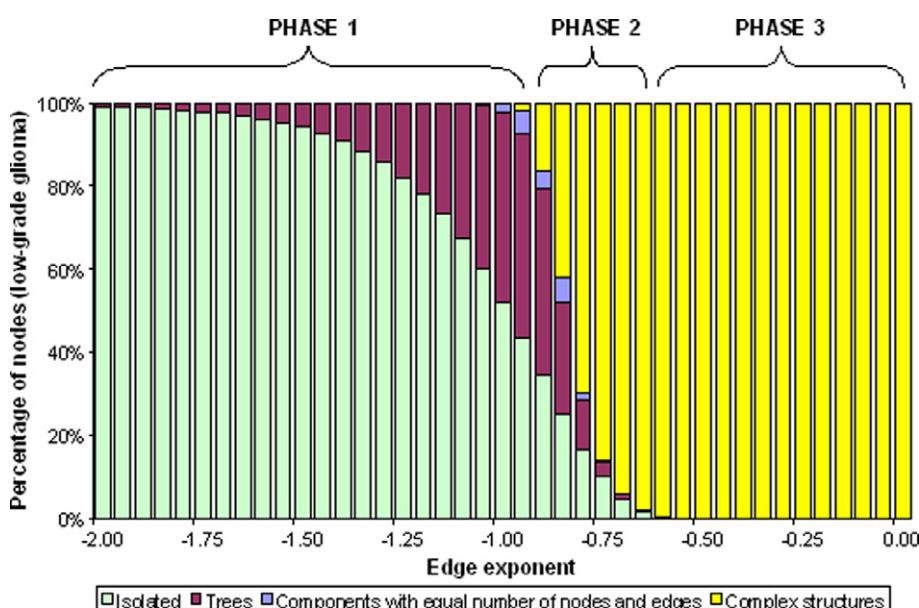
Fig. 6. For low-grade glioma, the distribution of nodes depending on the structure of the components that they belong to.
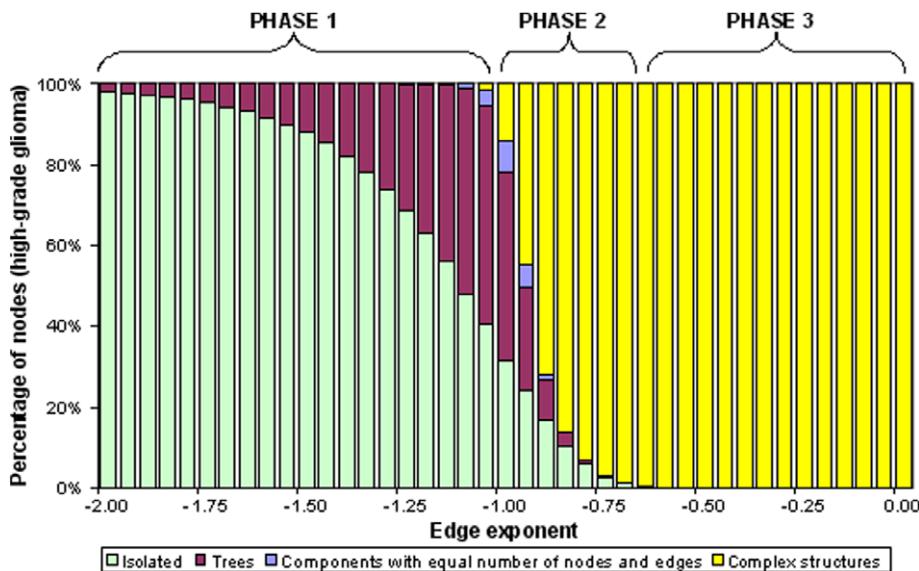
Fig. 7. For high-grade glioma, the distribution of nodes depending on the structure of the components that they belong to.

there exists significant difference between these distributions for the edge exponents smaller than −0.50. Regardless of the tissue type, edge exponents greater than −0.50 lead to connected graphs that have a single component with a complex structure.

4. Discussion

Our experiments on normal, low-grade cancerous (gliomatous), and high-grade cancerous (gliomatous) tissues have demonstrated that cell-graphs of different tissue types evolve differently although they are generated using the same number of nodes (cells) and the same edge exponent. Next, we explore the use of this difference in modeling the malignancy of cancer. For that, we gradually increase the edge exponent so long as it yields cell-graphs of Phase 1 for normal tissues (until $\alpha \leqslant -0.90$). When $\alpha = -0.90$, the cell-graph of a normal tissue has 60.00 edges ($<N/2$) and its giant connected component has 15.19 nodes, on the average. However, the same edge exponent yields cell-graphs of Phase 2 for cancerous tissues, regardless of their malignancy grades. For both the low-grade and high-grade glioma, the number of edges exceeds the $N/2$ threshold and the size of giant components exceeds the $N^{2/3}$ threshold. Moreover, this edge exponent leads to structural changes in the cell-graphs of both of these tissue types. On the other hand, the number of edges for the low-grade glioma (83.38 on the average) is much closer to the $N/2$ threshold compared to that of the high-grade glioma (136.45 on the average). Similarly, the giant connected component size for the low-grade glioma (39.95 on the average) is much closer to the $N^{2/3}$ threshold compared to that of the high-grade glioma (110.44 on the average). Additionally, there are more structural changes in the cell-graphs of the high-grade glioma compared to the case of the low-grade glioma.

These results demonstrate that the existence of cancer in a tissue transforms its cell-graph from Phase 1 to Phase 2 even though cell-graphs of normal tissues remain in Phase 1 with the same edge exponent. Therefore, this implies that cancer causes changes in the cell distribution of tissues, which is consistent with the common medical knowledge [25]. Furthermore, it is also known that these structural changes are aggravated with the increasing malignancy grade. Our results are also consistent with this such that cell-graphs of low-grade cancerous (gliomatous) tissues are at the boundary between Phase 1 and Phase 2 (just at the beginning of Phase 2) whereas cell-graphs of high-grade cancerous (gliomatous) tissues are more derived from those of normal tissues compared to those of low-grade ones.

We also test our algorithm on 24 different tissue images (8 images for each tissue type). These images are taken from 12 different patients where six of them are those that we have previously used (dependent patient group) and the remaining six are completely different patients (independent patient group). For each tissue image, we construct ten series of cell-graphs each of which is generated using the same edge exponents. For each tissue, we determine the phase using these ten instances of cell-graphs when $\alpha = -0.90$. To do so, we compare the number of edges and the giant connected component size of these ten cell-graphs against the specified thresholds ($N/2$ and $N^{2/3}$, respectively). A t-test with a significance level of 0.05 demonstrates that the corresponding values of normal tissues are statistically less than the specified thresholds. The t-test also shows that there are only two false negatives; there are two low-grade cancerous tissues with the corresponding values less than the specified thresholds. Note that one of these false negatives belongs to the dependent patient group whereas the other one belongs to the independent patient group. In the differentiation of cancerous tissues, we employ the mean values that we find for the low-grade and high-grade glioma. For each cancerous tissue, we compare the distance between the computed values and the means for the low-grade and high-grade glioma. A t-test with a significance level of 0.05 shows that for low-grade cancerous tissues, the computed values are statistically closer to the mean for the low-grade glioma than that for the high-grade glioma. Similarly, it shows that for high-grade cancerous tissues, the computed values are statistically closer to the mean for the high-grade glioma than that for the low-grade glioma.

Additionally, we compare our results against those obtained by the textural approach. In literature, a number of studies have proposed to use textural features, which provide information about the variation of an image surface, to represent a tissue image [26–29]. One of the most common methods to describe textural features is to use the co-occurrence matrix [30]. The co-occurrence matrix C is defined by a distance and an angle where $C(i,j)$ indicates how many times the intensity value i co-occurs with the intensity value j in a particular spatial relationship defined by the distance and the angle. Esgiar et al. derive six textural features including the angular second moment, contrast, correlation, inverse difference moment, dissimilarity, and entropy from multiple normalized gray-level co-occurrence matrices and use the average value of each feature to represent a tissue image [28].

Similarly, in our experiments, we derive the same textural features from normalized gray-level co-occurrence matrices, at four different angles (0°, 45°, 90°, and 135°) and three different distances (1, 5, and 9). The nearest neighbor algorithm that uses these textural features misclassifies four test samples two of which are from the independent patient group.[5] Here we observe that our

---

[5] In the nearest neighbor algorithm, we normalize the data for the distance computation.

approach leads to better results in this particular application. Additionally, one of the important features of our approach is the ability of demonstrating that the evolution process of cell-graphs is different for normal and cancerous (both low- and high-grade) tissues. The difference in the graph evolution process opens up new possibilities for dynamic modeling of cancer, which is one of the future research directions of our work. To the best of our knowledge, such ability of the textural approach has not been demonstrated for dynamic cancer modeling.

This work proposes a new method that uses graph evolution process to model the malignancy (grade) of cancer, for the first time. In this work, we present how to quantify graph evolution process through graph phases and how to use this phase information for the purpose of cancer grading. Our experiments on the brain tissue samples demonstrate that (1) cell-graphs of different tissue types evolve differently and (2) this difference can be used to model the malignancy grades of cancer, i.e., to identify the cell-graphs of normal, low-grade cancerous (gliomatous) and high-grade cancerous (gliomatous) tissues. One of the future research directions is to generate synthetic cell-graphs and to simulate cancer formation and progression through the use of graph phases. Another research opportunity is to investigate the use of this method in modeling the malignancy of different types of cancer.

Acknowledgments

References

[1] D. Watts, S. Strogatz, Collective dynamics of small-world networks, Nature 363 (1998) 202.
[2] M. Faloutsos, P. Faloutsos, C. Faloutsos, On power-law relationships of the Internet topology, in: Proceedings of ACM/SIGCOMM, Cambridge, Massachusetts, 1999, p. 251.
[3] R. Albert, H. Jeong, A.-L. Barabasi, Diameter of the World Wide Web, Nature 401 (1999) 130.
[4] A. Broder, R. Kumar, F. Maghoul, P. Raghavan, R. Stata, Graph structure in the Web, in: Proceedings of the ninth International World Wide Web Conference, Amsterdam, The Netherlands, 2000, pp. 247.
[5] M.E.J. Newman, Who is the best connected scientist? A study of scientific coauthorship networks, Phys. Rev. E64 (2001), cond–mat/0011144.
[6] L.A. Meyers, B. Pourbohloul, M.E. J Newman, D.M. Skowronski, R.C. Brunham, Network theory and SARS: predicting outbreak diversity, J. Theor. Biol. 232 (2005) 71.
[7] F. Liljeros, C.R. Edling, L.A.N. Amaral, H.E. Stanley, Y. Aberg, The web of human sexual contacts, Nature 411 (2001) 907.
[8] S. Milgram, The small-world problem, Psychol. Today 2 (1967) 61.
[9] S. Wasserman, K. Faust, Social Network Analysis: Methods and Applications, Cambridge University Press, Cambridge, UK, 1994.
[10] H. Jeong, S.P. Mason, A.-L. Barabasi, Z.N. Oltvai, Lethality and centrality in protein networks, Nature 411 (2001) 41.
[11] H. Jeong, B. Tombor, R. Albert, Z.N. Oltvai, A.-L. Barabasi, The large-scale organization of metabolic networks, Nature 407 (2000) 651.
[12] A. Wagner, D. Fell, The small world inside large metabolic networks, P. Roy. Soc. Lond. B Bio. 268 (2001) 1803.
[13] S. Wuchty, E. Ravasz, A.-L. Barabasi, The architecture of biological networks, in: T.S. Deisboeck, J. Yasha Kresh, T.B. Kepler (Eds.), Complex Systems in Biomedicine, Kluwer Academic Publishing, New York, 2003.

[14] C. Gunduz, B. Yener, S.H. Gultekin, The cell graphs of cancer, Bioinformatics 20 (2004) i145.

[15] C. Demir, S.H. Gultekin, B. Yener, Learning the topological properties of brain tumors, IEEE/ACM Trans. Comput. Biol. Bioinformatics 2 (2005) 262.

[16] C. Demir, S.H. Gultekin, B. Yener, Augmented cell-graphs for automated cancer diagnosis, Bioinformatics 21 (2005) ii7.

[17] G.D. Thomas, M.F. Dixon, N.C. Smeeton, N.S. Williams, Observer variation in the histological grading of rectal carcinoma, J. Clin. Pathol. 36 (1983) 385.

[18] A. Andrion, C. Magnani, P.G. Betta, A. Donna, F. Mollo, M. Scelsi, P. Bernardi, M. Botta, B. Terracini, Malignant mesothelioma of the pleura: interobserver variability, J. Clin. Pathol. 48 (1995) 856.

[19] N. Otsu, A threshold selection method from gray level histograms, IEEE Trans. Syst. Man Cyb. 9 (1979) 62.

[20] J. Serra, Image Analysis and Mathematical Morphology, Academic, London, UK, 1982.

[21] L. Vincent, P. Soille, Watersheds in digital spaces: an efficient algorithm based on immersion simulations, IEEE Trans. Pattern Anal. 13 (1991) 583.

[22] P. Erdos, A. Renyi, On the evolution of random graphs, Publ. Math. Inst. Hungarian Acad. Sci. 5 (1960) 17.

[23] P. Erdos, A. Renyi, On the strength of connectedness of a random graph, Acta Mathematica Scientia Hungary 12 (1961) 261.

[24] M. Karonski, A. Rucinski, The origins of the theory of random graphs, Mathematics of Paul Erdos; Graham, 1997, p. 311.

[25] P. Burger, B. Scheithauer, F.S. Vogel, Surgical Pathology of the Nervous System and Its Coverings, 160–241, Churchill Livingstone, New York, 2002.

[26] F. Schnorrenberg, C.S. Pattichis, C.N. Schizas, K. Kyriacou, M. Vassiliou, Computer-aided classification of breast cancer nuclei, Technol. Health Care 4 (1996) 147.

[27] P.W. Hamilton, P.H. Bartels, D. Thompson, N.H. Anderson, R. Montironi, Automated location of dysplastic fields in colorectal histology using image texture analysis, J. Pathol. 182 (1997) 68.

[28] A.N. Esgiar, R.N.G. Naguib, B.S. Sharif, M.K. Bennett, A. Murray, Microscopic image analysis for quantitative measurement and feature identification of normal and cancerous colonic mucosa, IEEE Trans. Inform. Technol. Biomed. 2 (1998) 197.

[29] D.K. Tasoulis, P. Spyridonos, N.G. Pavlidis, D. Cavouras, P. Ravazoula, G. Nikiforidis, M.N. Vrahatis, Urinary bladder tumor grade diagnosis using on-line trained neural networks, Lect. Notes Artif. Int. 2773 (2003) 199.

[30] R.M. Haralick, Statistical and structural approaches to texture, P. IEEE 67 (1979) 786.