ON THE CONTINUITY OF GRAPH PARAMETERS

M. HURSHMAN AND J. JANSSEN

Abstract. In this paper, we explore the mathematical properties of a distance function between graphs based on the maximum size of a common subgraph. The notion of distance between graphs has proven useful in many areas involving graph based structures such as chemistry, biology and pattern recognition. Graph distance has been used as a way of determining similarity between graphs. The distance function studied here forms a metric on isomorphism classes of graphs. We show that this metric induces the discrete topology. We also show that the distance between two graphs in the Erdös-Renyi probability space $G(n, p)$ almost always is near the maximum attainable value. Finally, we define a notion of continuity of graph parameters and relate it to a property of graphs that can be easily verified. We also determine whether the normalized versions of some common graph parameters are continous in this framework.

1. Introduction

The problem of comparing graphs has received a lot of attention over the past 30 years. In this paper, we will call the problem of comparing of graphs the *graph similarity problem*. In many other papers, the problem is called the graph matching problem [5] [8]. The term graph matching indicates that the main interest of the application is to find corresponding regions in graphs which match. In the graph similarity problem, the main interest is to assign an overall similarity score to indicate the level of similarity between two graphs. Any method used for graph matching can be turned into a method for graph similarity by assigning a score to the similarity of the two graphs. That is, the two problems are the same, only their focus is different. In graph similarity, we refer to the function d that assigns the similarity score as the *distance function*. This is because we think of the similarity score between two graphs as the distance between them. Though different similarity methods lead to different distance functions, we would like all distance functions to satisfy the three following properties.

- $d(G, H) = 0$ iff $G \cong H$
- $d(G, H) = d(H, G)$

- Increasing values of $d(G, H)$ indicate a decreasing amount of similarity between G and H.

The first condition states that if two graphs are isomorphic then the distance between them should be 0. The second requirement is that the distance function is symmetric. If the distance function additionally satisfies the triangle inequality, as is the case with the similarity method based on the maximum common subgraph [3], the distance function forms a metric on the set of isomorphism classes of graphs.

The main force driving the study of this problem has been its numerous applications in many areas such as biology, chemistry and pattern recognition [8] [15]. There have been many proposed methods such as graph and subgraph isomorphism, maximum common subgraph methods and edit distance [6]. More recently, the focus has shifted to the use of machine learning techniques and graph kernels [23]. The main interest in graph similarity has been to develop methods to determine graph similarity and efficient algorithms to compute the distance function. This has left many questions of a more mathematical flavour unexplored.

The purpose of this paper is to explore some of these questions using the distance metric of Bunke and Shearer based on the maximum common subgraph. The first contribution of this paper will be to show that the maximum common subgraph distance function induces the discrete topology on the set of isomorphism classes of graphs. We will also show that the distance between two graphs from the Erdös-Renyi probability space $G(n, p)$ almost always have a distance near 1, which is the maximum value. The implication is that two randomly chosen graphs are not likely to be similar. The main contribution of this paper will be to introduce the idea of continuity of graph parameters. There are many graph parameters which we commonly assign to graphs such as the diameter, chromatic number, number of vertices etc. We might expect that for some of these parameters, a small distance between graphs would indicate a small difference between their graph parameters as well. Such a notion has a flavour of continuity and in Section 5 we will define both a pointwise and a uniform version of graph parameter continuity. We will show that all graph parameters are trivially continuous with respect to the maximum common subgraph distance function in the pointwise sense and give a condition on which graph parameters are continuous in the uniform sense.

2. Graph Distance Metric

This metric introduced by Bunke and Shearer in [3] is based on the size of the largest common subgraph between two graphs.

Definition 2.1. *Given G, G_1 and G_2, we say that G is a common subgraph of G_1 and G_2 if there exist isomorphisms from G to a subgraph of G_1 and from G to a subgraph of G_2. A maximum common subgraph of G_1 and G_2 is a common subgraph of maximum size. We use the notation* $\mathbf{mcs(G_1, G_2)}$ *to denote the maximum size of any common subgraph of G_1 and G_2.*

It is important to note that a maximum common subgraph of two graphs is not necessarily unique. Note also that maximum common subgraphs do not need to be connected.

Definition 2.2. *The distance between two graphs $G_1, G_2 \in \mathcal{G}$ is defined as*

$$d_{mcs}(G_1, G_2) = 1 - \frac{mcs(G_1, G_2)}{max(|G_1|, |G_2|)} \tag{1}$$

where $max(|G_1|, |G_2|)$ is the maximum size of the graphs G_1 and G_2.

Clearly, d_{mcs} is symmetric; Bunke and Shearer showed in [3] that d_{mcs} also satisfies the triangle inequality. When $d_{mcs}(G_1, G_2) = 0$, it implies that $mcs(G_1, G_2) = |G_1| = |G_2|$, and thus G_1 and G_2 are isomorphic, but not necessarily equal. Thus, d_{mcs} is not a metric over all graphs, but it is a metric over all *isomorphism classes* of graphs. In the following, we will use $\mathcal{G}$ to denote the space of all isomorphism classes of non-empty graphs. We will use $\mathcal{G}_n$ to denote the set of isomorphism classes of graphs on n vertices. To avoid cumbersome notation, we will use the expression "$G \in \mathcal{G}$" to denote both the graph itself, and its isomorphism class.

3. $(\mathcal{G}, d_{mcs})$ Induces The Discrete Topology

As d_{mcs} forms a metric on the set $\mathcal{G}$, the pair $(d_{mcs}, \mathcal{G})$ forms a metric space. In this section we will show that this metric space is actually the discrete topology. We will show that d_{mcs} induces the discrete topology on $\mathcal{G}$ by showing that all open balls in the metric space are also open sets in the topology. We have the following definition for an open ball in $\mathcal{G}$.

Definition 3.1. *If $G \in \mathcal{G}$ and $r > 0$, then an open ball centered at G with radius r is the set*

$$B(G, r) = \{H \in G : d_{mcs}(G, H) < r\}$$

To consider what open balls look like, we consider what possible distances can exist between a fixed graph G and all other graphs in $\mathcal{G}$. By definition, $d_{mcs}(G, H) \in [0, 1) \cap \mathbb{Q}$. Since d_{mcs} is a metric, $d_{mcs}(G, H) = 0$ precisely when $G \cong H$. A distance of 1 is not possible as graphs in $\mathcal{G}$ are non-empty, so any two graphs G and H always have at least one vertex in common so that $mcs(G, H) \geq 1$. Therefore, for $r \geq 1$ we have that $B(G, r) = \mathcal{G}$.

Since $mcs(G, H)$ is an integer, for graphs G and H with sizes $|G| = n$ and $|H| = m \leq n$, $d_{mcs}(G, H) := \frac{n-i}{n}$ for some $i \in \{1, \ldots, m\}$. This observation leads to the following theorem.

Theorem 3.2. *Consider a fixed graph G of size n and let $H \in \mathcal{G}$. Then the minimum non zero value for $d_{mcs}(G, H)$ is $\frac{1}{n+1}$.*

Proof. Suppose G and H are not isomorphic, so $mcs(G, H) \leq \min\{|G|, |H|\}$. Let $|H| = m$. If $m \leq n$, the minimum distance between G and H with $H \in \mathcal{G}_m$ is $\frac{1}{n}$. If $m > n$, then the minimum distance between G and H with $H \in \mathcal{G}_m$ is $\frac{m-n}{m}$. We thus have to consider the minimum of the set $\{\frac{m-n}{m} : m > n\}$. The minimum distance occurs when $m = n+1$, so the minimum distance between G and H is then $d_{mcs}(G, H) = \frac{1}{n+1}$. □

Thus, if we take $r = \frac{1}{1+|G|}$ we get $B(G, r) = \{G\}$. Therefore, for all $G \in \mathcal{G}$, the set $\{G\}$ is open, which proves the following corollary.

Corollary 3.3. *The distance function d_{mcs} induces the discrete topology on $\mathcal{G}$.*

Theorem 3.2 says that for a fixed graph G, there is a limit on how small $d_{mcs}(G, H)$ can be for $H \in \mathcal{G}$. We can ask whether this is the case for any $G, H \in \mathcal{G}$. That is, can we find G, H so that $d_{mcs}(G, H)$ is as small as we like? This question is answered by the following theorem.

Theorem 3.4. *For any $q \in [0, 1) \cap \mathbb{Q}$, there exist two graphs G and H so that $d_{mcs}(G, H) = q$.*

Proof. Let $q = i/n$, where n and i are non-negative integers, and $i < n$. Let $G = K_n$. Let H be the disjoint union of the clique K_{n-i} and the set of isolated vertices $\overline{K_i}$. as follows. Then $mcs(G, H) = n - i$, and

$$d_{mcs}(G, H) = 1 - \frac{n-i}{n} = \frac{i}{n} = q$$

□

4. Distance between random pairs of graphs

In this section, we investigate the *mcs* distance between two randomly chosen graphs in the well-known Erdös-Renyi random graph probability space $G(n,p)$. The main theorem below states that almost any pair of randomly chosen graphs in $G(n,p)$ are at near-maximum distance from each other. Namely, the distance between these graphs tends to 1 as n tends to infinity, while the definition of *mcs* distance is such that this distance never can exceed 1. For the special case where $p = 1/2$, $G(n,p)$ gives the uniform distribution on all graphs with n vertices (all labelled graphs). Thus, a corollary of the the theorem is that almost all pairs of graphs of size n have *mcs* distance close to 1.

Theorem 4.1. *Let G and H be two graphs chosen according to $G(n,p)$. Then almost surely*

$$1 - \frac{5\log_{1/p^*}(n)}{n} < d_{mcs}(G,H) < 1 - \frac{2\log_{1/p^*}(n)}{n}.$$

Proof. Let G and H be two graphs chosen according to $G(n,p)$. Consider the graph GH formed as follows: $V(GH) = [n]$, and for each pair $i < j$ in $[n]$, i and j are adjacent in GH precisely when i and j are adjacent in both G and H, or i and j are non-adjacent in both G and H. Clearly, any clique of size k in GH corresponds to a common subgraph of size k in G and H. Also, the probability of an edge occurring in GH equals $p^* = p^2 + (1-p)^2 = 2p^2 - 2p + 1$, and edges still occur independently, so GH can be considered to be chosen according to $G(n,p^*)$. It a well known result in random graph theory (see for example [1]) that almost surely the clique number $G(n,p^*)$ is at least $2\log_{1/p^*} n$. This leads to upper bound on $d_{mcs}(G,H)$.

For the lower bound, we compute the probability that G and H have a common subgraph of size k. Let X be the number of common subgraphs of size k. For every set S, and every one-to-one map $f : S \to [n]$, define the indicator variable $X_{S,f}$, where $X_{S,f} = 1$ if f is an isomorphism from the subgraph of G induced by S to the subgraph of H induced by $f(S)$, and zero otherwise. Then $X = \sum_{S,f} X_{S,f}$, and by linearity of expectation, $\mathbb{E}(X) = \sum_{S,f} \mathbb{E}X_{S,f}$. For each particular choice of S and f,

$$\mathbb{E}X_{S,f} = \mathbb{P}(X_{S,f} = 1) = (p^*)^{\binom{k}{2}}.$$

Namely, for each pair $\{i,j\} \subseteq S$, either i,j are adjacent in G and $f(i), f(j)$ are adjacent in H, which happens with probability p^2, or i,j are non-adjacent in G and $f(i), f(j)$ are non-adjacent in H, with probability $(1-p)^2$.

There are $\binom{n}{k} \leq n^k$ choices for S, and, given S, there are $n(n-1)\ldots(n-k+1) \leq n^k$ choices for f. Therefore,

$$\mathbb{E}X \leq n^{2k}(p^*)^{\binom{k}{2}} \leq n^{2k}(1/p^*)^{-k^2/2} + o(1).$$

Now let $k = 5\log_{1/p^*}(n)$, so $(1/p^*)^{-k^2/2} = n^{-2.5k}$. Then $\mathbb{E}X = O(n^{-0.5k}) = o(1)$. By Markov's inequality,

$$\mathbb{P}(d_{mcs}(G,H) < \frac{n-k}{n}) = \mathbb{P}(X \geq 1) \leq \mathbb{E}X,$$

and $\mathbb{E}X \to 0$ when $n \to \infty$.

□

The reader may not think that this result extends to the unlabelled graphs of $\mathcal{G}$ since the proof assumes we are working with the labelled graphs of $G(n,p)$. This is not the case. For the upper bound, the maximum common subgraph between two labelled graphs is less than or equal to the maximum common subgraph between their corresponding unlabelled graphs so that the upper bound still holds. For the lower bound, we consider all subsets of vertices of both graphs and all one-to-one functions between them so that the lower bound given is precisely the lower bound for the unlabelled graphs as well.

5. Graph Continuity

In this section we introduce a notion of continuity for graph parameters. Our definitions follow the standard definition of continuity of functions, but are adapted to $(\mathcal{G}, d_{mcs})$.

Most commonly, continuity refers to *pointwise continuity*. In our setting, a graph function $f : \mathcal{G} \to \mathbb{R}$ is continuous at $G_0 \in \mathcal{G}$ if for all $\epsilon > 0$ there exists a $\delta > 0$ such that $d(G, G_0) < \delta$ implies that $|f(G) - f(G_0)| < \epsilon$. This definition is not very useful for the metric space under consideration, since all graph functions $f : \mathcal{G} \to \mathbb{R}$ are trivially pointwise continuous. Namely, by Theorem 3.2, if we select $\delta = \frac{1}{1+|G_0|}$, the only graph H satisfying $d_{mcs}(G, G_0) < \delta$ is $G = G_0$. Thus, every graph function f is pointwise continuous at G.

A more useful concept is that of uniform continuity.

Definition 5.1. *A graph function $f : \mathcal{G} \to \mathbb{R}$ is* uniformly continuous *if for all $\epsilon > 0$ and $G, H \in \mathcal{G}$ there exists a $\delta > 0$ such that when $d(G,H) < \delta$ we have that $|f(G) - f(H)| < \epsilon$.*

Many common graph functions, such as chromatic number, diameter, or girth, are integer valued. Thus, for such an integer valued function f, $f(G_1) - f(G_2)| < 1$ implies that $f(G_1) = f(G_2)$, making the condition for uniform continuity too restrictive. For this reason, we consider functions that are graph parameters normalized by the size of the graph.

Definition 5.2. *Let $f : \mathcal{G} \to \mathbb{R}$ be a graph function. The function $\bar{f}$ given by $\bar{f}(G) = \frac{f(G)}{|G|}$ is the normalized graph function for f.*

The remainder of this section explores the relation between a boundedness condition on a graph parameter, and uniform continuity of the derived normalized graph function.

Definition 5.3. *Given a graph parameter f, the step change of f equals the supremum, over all choices of graph G and vertex v of G, of $|f(G) - f(G - v)|$.*

Since the maximum is taken over the infinite family of all graphs, the step change can be infinite. Note that the step change is always non-negative.

Lemma 5.4. *Given a graph parameter with step change C. Then for any two graphs G and H of size $|H| \leq |G| = n$, $|f(G) - f(H)| \leq 2Cnd_{mcs}(G, H)$.*

Proof. Let $i = nd_{mcs}(G, H)$, and let K be a maximum common subgraph of G and H. Note that, by the definition of d_{mcs}, i is an integer, and $|K| = n - i$. Let $A = \{a_1, \dots, a_i\} \subseteq V(G)$ and $B = \{b_1 \dots, b_j\} \subseteq V(H)$ be so that $G - A$ and $H - B$ are isomorphic to K. By repeated application of the definition of the step change we obtain that $|f(G) - f(G - a_1 - \dots - a_i)| \leq iC$. Similarly, $|f(H) - f(H - b_1 - \dots - b_j)| \leq Cj$. Since $f(G - a_1 - \dots - a_i) = f(K) = f(H - b_1 - \dots - b_j)|$, we can conclude that $|f(G) - f(H)| \leq |f(G) - f(K)| + |f(H) - f(K)| \leq 2Ci$. □

Theorem 5.5. *Let f be a graph function that satisfies $f(G) \leq |G|$ for all G, and has finite step change C. Then $\bar{f}$ is uniformly continuous.*

Proof. Let f be a graph function and G and H are graphs with $|G| = n$ and $|H| = m$ and $n \geq m$. Let $i = nd_{mcs}(G, H)$. By the definition of the step change, this implies that $|f(G) - f(H)| \leq 2Ci$. Now let $\bar{f}$ be the normalized graph function for f.

Consider $|\bar{f}(G) - \bar{f}(H)|$ and write $f(G) = p_1$ and $f(H) = p_2$. We have that

$$\begin{aligned} |\bar{f}(G) - \bar{f}(H)| &= \left| \frac{p_1}{n} - \frac{p_2}{m} \right| \\ &= \left| \frac{mp_1 - np_2}{nm} \right| \\ &= \left| \frac{mp_1 - mp_2 - np_2 + mp_2}{mn} \right| \\ &\leq \left| \frac{p_1 - p_2}{n} \right| + \left| \frac{p_2(n-m)}{nm} \right| \end{aligned}$$

Now

$$|\frac{p_1 - p_2}{n}| = |\frac{f(G) - f(H)}{n}| \leq 2Ci/n = 2Cd_{mcs}(G, H).$$

Consider the second piece of our inequality $\frac{p_2(n-m)}{nm}$. By assumption, $p_2 = f(H) \leq |H| = m$, so $\frac{p_2}{m} \leq 1$ giving

$$\frac{p_2(n-m)}{nm} \leq \frac{n-m}{n} \leq \frac{n - |mcs(G, H)|}{n} = d_{mcs}(G, H).$$

Therefore,

$$|\bar{f}(G) - \bar{f}(H)| \leq (1 + 2C)d_{mcs}(G, H).$$

Fix $\epsilon > 0$, and let $\delta = \frac{\epsilon}{1+2C}$. Then $d_{mcs}(G, H) < \delta$ implies $|\bar{f}(G) - \bar{f}(H)| < \epsilon$. Thus, $\bar{f}$ is uniformly continuous.

□

The following theorem gives a result that is close to the converse of the above theorem.

Theorem 5.6. *Let f be a graph parameter so that $\bar{f}$ is uniformly continuous. Then for all graphs G and H where $|H| \leq |G| = n$ and $d_{mcs}(G, H) = \frac{i}{n}$,*

$$|f(G) - f(H)| \leq \alpha(n - i + 1)i,$$

Where $\alpha : \mathbb{N} \to \mathbb{N}$ is a function such that $\frac{\alpha(n)}{n} \to 0$ as $n \to \infty$.

Proof. Let

$$\alpha'(n) = max_{|G|=n,\, v \in V(G)} |\bar{f}(G) - \bar{f}(G - v)|,$$

We first show that $\alpha'(n)/n \to 0$ as $n \to \infty$. Fix $\epsilon > 0$. Since $\bar{f}$ is uniformly continuous, there exists $\delta > 0$ be such that $d_{mcs}(G,H)) < \delta$ implies that $|\bar{f}(G) - \bar{f}(H)| < \epsilon$.

Take $N \in \mathbb{N}$ such that $\frac{1}{N} < \delta$. Fix $G \in \mathcal{G}$ of size $n \geq N$, and $H = G - v$ for some vertex $v \in G$. Then $d_{mcs}(G,H) = 1/n \leq 1/N < \delta$. Thus we have that

$$|f(G) - f(H)| = |n\bar{f}(G) - (n-1)\bar{f}(H)| \leq n|\bar{f}(G) - \bar{f}(H)| \leq \epsilon n.$$

So for all $\epsilon > 0$, there exists $N \in \mathbb{N}$ so that $\alpha'(n)/n \leq \epsilon$ for all $n \geq N$.

Now let

$$\alpha(n) = n \sup\{\alpha'(k)/k \mid k \geq n\}.$$

Clearly, since $\alpha'(n)/n$ goes to zero as n goes to infinity, then so does $\alpha(n)/n$.

Fix G and H with maximum common subgraph K, and $|H| \leq |G| = n$. Let $i = nd_{mcs}(G,H)$. By the definition of the step change it follows that

$$\begin{aligned} |f(G) - f(H)| &\leq \alpha'(n) + \alpha'(n-1) + \cdots + \alpha'(n-i+1) \\ &\leq n\left(\tfrac{\alpha'(n)}{n} + \tfrac{\alpha'(n-1)}{n-1} + \cdots + \tfrac{\alpha(n-i+1)}{n-i+1}\right) \\ &\leq i\alpha(n-i+1) \end{aligned}$$

□

Theorem 5.5 makes it easy to find examples of graph parameters whose normalized functions are uniformly continuous. For example, the size of a graph, the chromatic number of a graph, and the domination number of a graph all have step change 1, and thus their normalized functions are uniformly continuous.

We conclude with an example of a graph function which is not continuous.

Theorem 5.7. *The diameter is not a uniformly continuous graph function with respect to d_{mcs}*

Proof. Consider the graphs P_n and C_n and let $\bar{f}$ be the normalized graph function for the diameter. For all n, we have that $d(P_n, C_n) = \frac{1}{n}$ and $\bar{f}(C_n) = \frac{\lfloor \frac{n}{2} \rfloor}{n}$ and $\bar{f}(P_n) = \frac{n-1}{n} = 1$. For all n the difference in the normalized diameters is then

$$|\bar{f}(C_n) - \bar{f}(P_n)| = \left| \frac{\lfloor \frac{n}{2} \rfloor - (n-1)}{n} \right| \geq \frac{1}{2} - \frac{1}{n}$$

for $n \geq 1$. Fix $\epsilon = \frac{1}{4}$. Fix any $\delta > 0$, and let $N \geq \max\{4, 1/(N+1))\}$. Then $d_{mcs}(P_N, C_N) = 1/N < \delta$, but

$$|\bar{f}(C_N) - \bar{f}(P_N)| \geq \frac{1}{2} - \frac{1}{N} \geq \frac{1}{4}.$$

So, for every choice of $\delta > 0$, there exists a pair of graphs that violates the uniformity condition. □

References

[1] N. Alon, J.H. Spencer, *The Probabilistic Method*, Wiley, Second Edition, 2000.

[2] A.T. Balaban, *Chemical Applications of Graph Theory*, Academic Press Inc, 1976.

[3] H. Bunke, K. Shearer, *A graph distance metric based on the maximum common subgraph*, Pattern Recognition Letters **19**, 255-259, 1998.

[4] H. Bunke, X. Jiang, A. Kandel, *On the Minimum Common Supergraph of two Graphs*, Computing 65, No.1, 13-25, 2000.

[5] H. Bunke, *Graph Matching: Theoretical Foundation, Algorithms, and Applications*, in Proc. Vision Interface 2000, Montreal, 82-88, 2000.

[6] H. Bunke, *Error Correcting Graph Matching: On the Influence of the Underlying Cost Function*, IEEE Trans. Pattern Analysis and Machine Intelligence, Vol 21, No.1, 917-922, 1999.

[7] H. Bunke, *On a relation between graph edit distance and maximum common subgraph*, Pattern Recognition Letters, **18**, 689-694, 1997.

[8] D. Conte, P.Foggia, C. Sansone and M.Vento, *Thirty Years of Graph Matching in Pattern Recognition*, International Journal of Pattern Recognition and Artificial Intelligence, Vol. 18, No.3, 265-298, 2004.

[9] R. Diestel, *Graph Theory*, Springer, Third Edition, 2006.

[10] M.L. Fernández, G. Valiente *A graph distance metric combining maximum common subgraph and minimum common supergraph*, Pattern Recognition Letters, **22**, 753-758, 2001.

[11] M.R. Garey, D.S. Johnson, *Computers and Intractability: A Guide to the Theory of NP-Completeness*, Series of Books in Mathematics Sciences, 1979.

[12] G. Grimmett, D. Stirzaker, *Probability and Random Processes*, Prentice Hall, 3rd Edition, 2007.

[13] S. Janson, T. Łuczak, A. Rucinski, *Random Graphs*, Wiley, First Edition, 2000.

[14] I. Koch, *Enumerating all connected maximum subgraphs in two graphs*, Theoretical Computer Science, 1-30, 2001.

[15] S.Q. Le, T.B. Ho, T.T.H. Phan, *A Novel Graph-Based Similarity Measure for 2D Chemical Structures*, Genome Informatics, **15**, 82-91, 2004.

[16] B.D. McKay, *Practical graph isomorphism*, Congressus Numerantium 30: 4587, 10th. Manitoba Conference on Numerical Mathematics and Computing (Winnipeg, 1980)

[17] J.R. Munkres, *Topology*, Prentice Hall, 2nd Edition, 2000.

[18] E.M. Palmer, *Graphical Evolution: An Introduction to the Theory of Random Graphs*, Wiley, First Edition, 1985.

[19] P. Papadimitriou, A. Dasdan, H. Garcia-Molina, *Web Graph Similarity for Anomaly Detection* WWW 2008, Beijing, China, April 21-25 2008.

[20] J.W. Raymond, P. Willet *Maximum Common Subgraph Isomorphism Algorithms for the Matching of Chemical Structures*, Journal of Computer-Aided Molecular Design, 521-533, 2002.
[21] W.D. Wallis, P. Shoubridge, M. Kraetz, D. Ray, *Graph distances using graph union*, Pattern Recognition Letters, **22**, 701-704, 2001.
[22] D.B West, *Introduction to Graph Theory*, Prentice Hall, 2nd Edition, 2001.
[23] S.V.N Vishwanathan, K.M. Bogwardt and N.N. Schraudolph, *Fast computation of Graph Kernels*, Advances i Neural Information Processing Systems 19 (NIPS 2006), MIT Press, cambridge, MA, 2006.

Department of Mathematics and Statistics, Dalhousie University, Halifax, NS B3H 3J5, Canada