Electric Network Kernel for Support Vector Machines

01507724 Kyoto University HIRAI Hiroshi*
01603194 University of Tokyo MUROTA Kazuo
Justsystem RIKITOKU Masaki

1 Introduction

Support vector machine (SVM) [3] has come to be very popular in machine learning and data mining communities. Input data from real world problems is often endowed with discrete structures. Recently Kondor and Lafferty [2] introduced *diffusion kernels*, which are discrete kernels defined on vertices of graphs.

We propose a novel class of discrete kernels, named *electric network kernel*, on vertices of an undirected graph. SVM with this kernel admits physical interpretations in terms of resistive electric networks; in particular, the SVM decision function corresponds to an electric potential. Preliminary computational results are also reported.

2 Support Vector Machines

Let $\mathcal{X}$ be an input data space and $K : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ be a *kernel* on $\mathcal{X}$. Given a labeled training set $\{(x_i, \eta_i)\}_{i=1,\ldots,m} \subseteq \mathcal{X} \times \{\pm 1\}$, SVM classifier is obtained by solving the optimization problem [SVM]

$$\begin{aligned} \min\ & \frac{1}{2}\sum_{i,j=1}^{m} u_i u_j K(x_i, x_j) - \sum_{i=1}^{m} \eta_i u_i \\ \text{s.t.}\ & \sum_{i=1}^{m} u_i = 0,\ 0 \le \eta_i u_i \le C \quad (i = 1, \ldots, m), \end{aligned}$$

where C is a penalty parameter that is a positive real number or $+\infty$. Let $u^* \in \mathbb{R}^m$ be an optimal solution of the problem [SVM] and $b^* \in \mathbb{R}$ be the Lagrange multiplier of the equality constraint at u^*. Then the *decision function* $f : \mathcal{X} \to \mathbb{R}$ is given as

$$f(x) = \sum_{i=1}^{m} u_i^* K(x_i, x) + b^* \quad (x \in \mathcal{X}). \qquad (1)$$

We classify a given data x according to the sign of $f(x)$. A data x_i with $\eta_i u_i^* > 0$ is called a *support vector*.

3 Electric Network Kernel

Let (V, E, r) be a resistive electric network with vertex set V, edge set E, and resistors on edges with the resistances represented by $r : E \to \mathbb{R}_{>0}$. Let D be a distance function on V defined as

$$D(x, y) = \text{resistance between } x \text{ and } y$$

for $x, y \in V$. Fix some vertex $x_0 \in V$ as a root, and define an *electric network kernel* K on V as

$$K(x, y) = \{D(x, x_0) + D(y, x_0) - D(x, y)\}/2$$

for $x, y \in V$.

We give physical interpretations to the problem [SVM] on (V, E, r) with the aid of nonlinear network theory (see [1, Chapter IV]). Suppose that we are given an electric network (V, E, r) and labeled training data set $\{(x_i, \eta_i)\}_{i=1,\ldots,m} \subseteq V \times \{\pm 1\}$. We connect voltage sources to (V, E, r) as follows:

> For each x_i with $1 \le i \le m$, connect to the earth a voltage source whose electric potential is η_i and the current flowing into x_i is restricted to $[0, C]$ if $\eta_i = 1$ and $[-C, 0]$ if $\eta_i = -1$.

By using voltage sources, current sources and diodes, this network can be realized as in Figure 1.

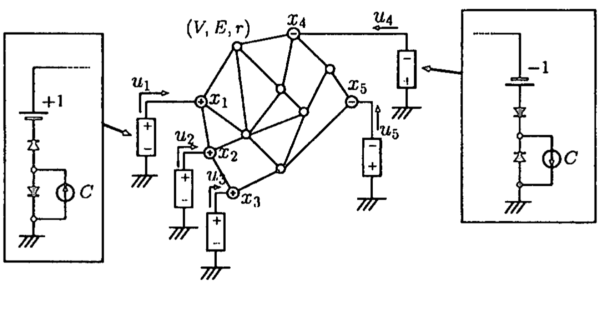

Figure 1: Physical interpretation

The following theorem indicates the relationship between SVM problem and this electric network.

Theorem 1. *Let u^* be the optimal solution of* [SVM] *on* (V, E, r). *Then u_i^* coincides with the electric current flowing into x_i for $i = 1, \ldots, m$. Moreover, the decision function f of* (1) *for* [SVM] *is an electric potential.*

Hence, the following correspondence holds.

SVM	electric network
positive label data	+1 voltage sources
negative label data	−1 voltage sources
optimal solution	current
decision function	potential

4 SVM on Tensor Product of Complete Graphs

We consider the case where (V, E) is an N-tensor product of k-complete graphs defined as

$$\begin{aligned} V &= \{0, 1, 2, \ldots, k-1\}^N, \\ E &= \{xy \mid x, y \in V,\ d_{\mathrm{H}}(x, y) = 1\}, \end{aligned}$$

where $d_{\mathrm{H}} : V \times V \to \mathbf{R}$ is the Hamming distance defined as

$$d_{\mathrm{H}}(x, y) = \{i \in \{1, \ldots, N\} \mid x^i \neq y^i\},$$

where x^i denotes the ith component of $x \in V$. Exploiting symmetry of this graph, the resistance D between two vertex pair is given as follows.

Theorem 2. *The resistance D of an N-tensor product of k-complete graphs (V, E) is given by*

$$D(x, y) = \begin{cases} \dfrac{1}{2^{N-2}} \displaystyle\sum_{s=1,3,5,\ldots}^{d} \sum_{u=0}^{N-d} \binom{d}{s} \binom{N-d}{u} \frac{1}{2(s+u)} & \text{if } k = 2, \\ \displaystyle\sum_{s=1,3,5,\ldots}^{d} \sum_{t=0}^{d-s} \sum_{u=0}^{N-d} \binom{d}{s} \binom{d-s}{t} \binom{N-d}{u} \times \frac{2^{2-s} k^{-N+s+t+u}}{k(s+t+u)} \left(\frac{1}{2} - \frac{1}{k}\right)^t \left(1 - \frac{1}{k}\right)^u & \text{if } k \geq 3, \end{cases}$$

where $d = d_{\mathrm{H}}(x, y)$.

The theorem implies, in particular, that each element of kernel K can be computed with $O(N^4)$ arithmetic operations. This makes it possible to apply the electric network kernel to large-scale practical problems on this class of graphs.

5 Experimental Results

Here, we describe preliminary experiments with our electric network kernels on tensor products of complete graphs. Table 1 shows the experimental results with Hamming kernel (HK), diffusion kernel (DK), and electric network kernel (ENK) for benchmark data sets taken from UCI Machine Learning Repository, where Acc means the ratio of correct answers and SVs is the number of support vectors.

The above results indicate that our electric network kernel works well as an SVM kernel. However comprehensive computational study is left as a future research topic.

Table 1: Experimental results

	HK		DK		ENK	
Data set	SVs	Acc	SVs	Acc	SVs	Acc
Hepat	60	79.1	60	79.8	106	77.7
Votes	36	96.0	53	96.0	274	84.5
LED2-3	386	89.6	392	89.7	388	89.8
Cancer	152	97.3	242	97.0	463	81.7

References

[1] M. Iri: *Network Flow, Transportation and Scheduling*, Academic Press, New York, 1969.

[2] R. I. Kondor and J. Lafferty: Diffusion kernels on graphs and other discrete structures, in: *Proceedings of the 19th ICML*, Morgan Kaufmann, 2002, 315–322.

[3] V. Vapnik: *Statistical Leaning Theory*, Wiley, 1998.