

## Random Walk on a Graph

It is well-known that *random walk on a graph* is one of fundamental research topics in the field of discrete mathematics. In the literature, a variety of studies have been devoted to understand the property of random walk on a graph through mathematical analysis and simulation experiment.

The rationale behind this situation is mainly caused by favorable property of random walk, in particular, applicability and simplicity. As will be described later, random walk on a graph means that a walker sequentially moves according to the structure of graph, which does not require any complex mechanisms. From the viewpoint of computer science, random walk on a graph is a promising approach to solve various problems for graphs. Because the graph can be regarded as an abstraction of *networks* such as communication network and social network, random walk on a graph is applicable to, for instance, content discovery on P2P (Peer-to-Peer) network and graph sampling on unknown social network.

Before describing random walk on a graph in detail, we provide a brief review on graph. Graph is one of typical data structures, and it is comprised of two fundamental components: *vertex* and *edge*. Edge is corresponding to a line between two distinct vertices, in other words, edge  $e$  is expressed as  $e = \{u, v\}$ . Formally, graph  $G$  is represented as  $G = (V, E)$ , where  $V$  and  $E$  represent a set of vertices and a set of edges, respectively. Also, a graph can be categorized into *directed* graph and *undirected* graph, according to orientation of an edge. In the following, we focus on the undirected graph rather than the directed graph for its simplicity.

Random walk on a graph is a series of movements that an *agent*, i.e., walker, traverses a graph. Specifically, the operation of an agent is described as follows; (i) an agent on vertex  $u$  randomly selects vertex  $v$  of vertices that are adjacent to vertex  $u$ , and the agent moves to the randomly-selected vertex  $v$  among adjacent vertices; (ii) procedure (i) is repeated until a given condition is satisfied. Strictly speaking, at step  $t$ , an agent on vertex  $u$  selects vertex  $v \in \mathcal{N}(u)$  of a set of adjacent vertices  $\mathcal{N}(u)$  which is defined as  $\{v \mid \{u, v\} \in E\}$ ; then, at step  $t + 1$ , an agent visits adjacent vertex  $v$ . Let us mathematically consider the behavior that the agent transits to a randomly-chosen adjacent vertex. We denote the probability that an agent on vertex  $u$  transits to vertex  $v$ , i.e., the transition probability, as  $p_{uv}$ . Also, we can simply formulate the transition probability as

$$p_{uv} = \frac{1}{d_u}, \quad (1)$$

where  $d_v$  represents the *degree* of vertex  $v$  and it is defined as the number of adjacent vertices of vertex  $v$ . Note that this definition is valid in the case of simple undirected graph.

The goodness of random walk on a graph can be evaluated with typical measures — *first hitting time* and *cover time*. Below, we provide definitions of these two measures. First, the first hitting time  $H_{st}$  is defined as the expected number of steps that an agent initially visits destination vertex  $t$  from source vertex  $s$ . Hence, the average first hitting time is given by the mean of the first hitting time  $H_{st}$  of every vertices pair  $(s, t)$  on graph  $G$ . Second, the cover time is defined as the expected number of steps that an agent reaches all vertices on a given graph from an arbitrary source vertex.