Tohoku Math. J.
63 (2011), 665–695

WEB MARKOV SKELETON PROCESSES AND THEIR APPLICATIONS

YUTING LIU, ZHI-MING MA AND CHUAN ZHOU

(Received December 1, 2010, revised May 9, 2011)

Abstract. We propose and discuss a new class of processes, web Markov skeleton processes (WMSP), arising from the information retrieval on the Web. The framework of WMSP covers various known classes of processes, and it contains also important new classes of processes. We explore the definition, the scope and the time homogeneity of WMSPs, and discuss in detail a new class of processes, mirror semi-Markov processes. In the last section we briefly review some applications of WMSPs in computing page importance on the Web.

1. Introduction. In this paper we propose and discuss a new class of processes, web Markov skeleton processes (WMSP), arising from the information retrieval on the Web.

Intuitively, a Markov skeleton process is a stochastic process which contains a Markov chain as its skeleton. A *web Markov skeleton process* is a jump process and also a Markov skeleton process such that, given the information of its skeleton, the time slots between jumps are conditionally independent to each other. The dynamics of a WMSP can be described as follows:

$$X_0 \xrightarrow{Y_0} X_1 \xrightarrow{Y_1} \cdots X_n \xrightarrow{Y_n} \cdots$$

where $\{X_n, n \geq 0\}$ is a Markov chain, and $\{Y_n, n \geq 0\}$ is the set of time slots between jumps. The length of each Y_n may depend on the states $\{X_n\}$, but will be independent of other $Y_k, k \neq n$, when the information of $\{X_n\}$ is known.

The dynamics described above appears in various natural and social sciences, such as biology, finance, queueing theory, engineering. A direct motivation of investigating such dynamics comes from the information retrieval on the Web. The notion of WMSP was recently invented in [9, 10], in which the authors found that WMSP is a very suitable framework for modeling the user browsing behavior on the Web. When modeling the user browsing behavior, the state space E is a collection of web pages, $X = \{X_n, n \geq 0\}$ describes the transition behavior between pages, which forms a Markov chain with state space E, and $Y = \{Y_n, n \geq 0\}$ represents staying times on the pages. The staying time on the current page is a random variable which may depend on the information of the current page and some other pages the user has visited or will visit, but in general will not depend on the other staying times. The framework of WMSP is very useful for computing web page importance, it provides us a unified

---

2000 *Mathematics Subject Classification.* Primary: 60J75; Secondary: 60G55, 60K15, 68P20.

Key words and phrases. Web Markov skeleton process, semi-Markov process, mirror semi-Markov process, information retrieval on the Web.

mathematical instrument applicable to most known algorithms for ranking web pages and websites, such as PageRank [4, 18, 27], AggregateRank [1, 2, 8], TrustRank [12], Block-level PageRank [5], PopRank [26], BrowseRank [19, 20, 21], and so on. Moreover, the framework is essential in designing some new algorithms to handle more complex problems. For example, mirror semi-Markov processes, a new class of processes in WMSP family, plays an essential role in designing MobileRank for computing page importance of the mobile Web [9, 10]. It is known that the structure of the mobile Web differs a lot from the usual Internet Web [17].

Theoretically, the framework of WMSP covers various known classes of processes, including discrete time Markov chains, time homogeneous continuous time Markov processes (Q-processes), semi-Markov processes, and others. It contains also important new classes of processes, such as simple WMSPs, which are of importance in theoretical study, and mirror semi-Markov processes, which are important in modeling browsing behavior on the mobile Web. Because of its natural and trackable structure, we expect that in the future there will be more new classes of processes to be found useful within the framework of WMSP.

In this paper, we explore for the first time the theoretical aspects of WMSP. In the next section we give a rigorous mathematical definition of WMSP. Along with it we clarify that our notion of Markov skeleton processes is more general than that which was previously proposed by Hou et al in [13, 14], because we need to meet the demand of information retrieval on the Web. In Section 3 we describe the scope of WMSP. Among other things, we include a complete proof that the semi-Markov processes previously discussed in the literature (which we refer to as classical semi-Markov processes) are special cases of our semi-Markov processes within the framework of WMSP. Section 4 is devoted to describing the time homogeneity of WMSPs. The idea employed in this section might be useful elsewhere. In Section 5 we discuss in detail a new class of processes, mirror semi-Markov processes, which is essential in modeling the browsing behavior on the mobile Web. In Section 6 we provide two results of limit distributions, which will be used in computing page importance on the Web. The remark of Section 7 suggests that the class of simple WMSPs is of importance in theoretical study. Finally, in Section 8 we briefly review some applications of WMSPs to the web page ranking. This section might be interesting for those readers who are not familiar with information retrieval but who want to know how mathematics can be well used in the study of information retrieval on the Web.

Before we conclude this introduction, we would like to remark that the study of the theory of WMSPs is just beginning. Considering the length of this paper, some important topics, such as the theory of multivariate point process (cf. [16]), martingale methods (cf. [15, 24]), stability and reconstruction problems ([23]), will be discussed in our subsequent papers. We hope that this paper will stimulate more researches on this important and new class of stochastic processes.

2. Definition of WMSP.

In this section we introduce the notion of *web Markov skeleton process* (WMSP in short). We start with a general definition of *Markov skeleton processes*.

Let $(\Omega, \mathscr{F}, P)$ be a complete probability space. Let E be a metrizable Lusin space. We adjoint an extra point Δ (the cemetery) to E as an isolated point. Write $\tilde{E}$ for $E \cup \{\Delta\}$ and $\mathscr{E}$ for the Borel sets of $\tilde{E}$. We consider a right continuous stochastic process $Z = \{Z_t, t \geq 0\}$ with state space E and life time ζ. That is, Z_t is a $\tilde{E}$ valued random variable on Ω for each $t \geq 0$ and $Z_\cdot(\omega)$ is right continuous for each $\omega \in \Omega$, ζ is a $(0, \infty]$ valued random variable, $Z_t(\omega) \in E$ for $t < \zeta(\omega)$, and $Z_t(\omega) = \Delta$ for $t \geq \zeta(\omega)$. Let $(\mathscr{F}_t)_{t\geq 0}$ be the natural filtration generated by Z. Namely, $(\mathscr{F}_t)_{t\geq 0}$ is the minimal right continuous increasing family of sub-σ-algebras of $\mathscr{F}$ such that $\mathscr{F}_0$ contains all P-null sets and Z_t is $\mathscr{F}_t$ measurable for each t. Then $(\Omega, \mathscr{F}, (\mathscr{F}_t), P)$ forms a filtration space, Z is an $(\mathscr{F}_t)$ adapted process and ζ is an $(\mathscr{F}_t)$ stopping time. Recall that a random variable $\tau : \Omega \to [0, \infty]$ is called an $(\mathscr{F}_t)$ stopping time (or simply stopping time) if $\{\tau \leq t\} \in \mathscr{F}_t$ for all $t \geq 0$. For a stopping time τ, the σ-algebra prior to τ is defined by $\mathscr{F}_\tau := \{A \in \mathscr{F}; A \cap \{\tau \leq t\} \in \mathscr{F}_t \text{ for all } t \geq 0\}$.

Intuitively, a Markov skeleton process (MSP) is a stochastic process which contains a Markov chain as its skeleton. We propose the definition of MSP as follows. For notational convenience, we shall sometimes write $Z(t)$ for Z_t.

DEFINITION 2.1. Let $Z = \{Z(t), t \geq 0\}$ be a right continuous stochastic process with state space E and life time ζ. With the above notations, we say that Z is a *Markov skeleton process* (MSP in short) if there exists a sequence of stopping times $\{\tau_n\}_{n\geq 0}$ such that $0 = \tau_0 < \tau_1 < \cdots < \tau_n < \tau_{(n+1)} < \cdots < \zeta$, $\lim_{n\to\infty} \tau_n = \zeta$, and $\{Z(\tau_n)\}_{n\geq 0}$ forms a Markov chain, i.e., for all $n \geq 0$, $m \geq 1$, and $B_k \in \mathscr{E}$, $1 \leq k \leq m$, it holds that

$$(1) \quad P(Z_{\tau_{(n+k)}} \in B_k, 1 \leq k \leq m \mid Z_{\tau_0}, Z_{\tau_1}, \ldots, Z_{\tau_n}) = P(Z_{\tau_{(n+k)}} \in B_k, 1 \leq k \leq m \mid Z_{\tau_n}).$$

$\{Z(\tau_n)\}_{n\geq 0}$ will be referred to as the *Markov skeleton* or the *embedded Markov chain* of Z.

We remark that in the literature the notion of Markov skeleton process has been introduced by Hou el al. in [13, 14], and the authors of [13, 14] studied their MSPs intensively. But the framework of [13, 14] is not applicable to some applications in the research of Internet information retrieval. Therefore we propose our Definition 2.1 of MSP, which is more general than the one that proposed in [13, 14]. To simplify the terminology, in this paper we shall call a Markov skeleton process in the sense of [13, 14] as Hou's MSP. To compare the two notions, we restate the definition proposed in [13] as follows.

DEFINITION 2.2 (cf. [13, Definition 1]). Let $Z = \{Z(t), t \geq 0\}$ be a right continuous stochastic process with state space E and life time ζ. With the above notations, we say that Z is *Hou's Markov skeleton process* (Hou's MSP in short) if there exists a sequence of stopping times $\{\tau_n\}_{n\geq 0}$ satisfying $0 = \tau_0 < \tau_1 < \cdots < \tau_n < \tau_{(n+1)} < \cdots < \zeta$, $\lim_{n\to\infty} \tau_n = \zeta$, such that for all $n \geq 0$ and all bounded $\mathscr{E}^{[0,\infty)}$ measurable functions f defined on $\tilde{E}^{[0,\infty)}$, it holds that

$$(2) \qquad E\big[f\big(Z(\tau_n + \cdot)\big)\big|\, \mathscr{F}_{\tau_n}\big] = E\big[f\big(Z(\tau_n + \cdot)\big)\big|\, Z_{\tau_n}\big].$$

Comparing (1) and (2), it is easy to verify that Hou's MSP is always a Markov skeleton process in the sense of Definition 2.1. But the converse is not true. For example, a mirror semi-Markov process, which will be discussed in detail in Section 5 below, is a MSP in the sense of Definition 2.1, but not a Hou's MSP.

Let $Z = \{Z(t), t \geq 0\}$ be a Markov skeleton process. We use the notations $X_n = Z(\tau_n)$, and $Y_n = \tau_{n+1} - \tau_n$, for all $n = 0, 1, \ldots$. Then $X = \{X_n, n \geq 0\}$ is a Markov chain and $Y = \{Y_n, n \geq 0\}$ is a sequence of positive random variables. Clearly the pair (X, Y) is uniquely determined by the MSP Z. Conversely, if

$$Z(t) = X_n \quad \text{for } \tau_n \leq t < \tau_{n+1}\,, \quad \text{for all } n \geq 0\,, \tag{3}$$

then Z is also uniquely determined by (X, Y). We shall say that a Markov skeleton process Z is a *Markov Skeleton Jump Process* if it satisfies (3).

Suppose that we are given a Markov chain $X = \{X_n, n \geq 0\}$ and a sequence of positive random variables $Y = \{Y_n, n \geq 0\}$. Then a Markov skeleton jump process Z can be uniquely determined by the following regime:

$$X_0 \xrightarrow{Y_0} X_1 \xrightarrow{Y_1} \cdots X_n \xrightarrow{Y_n} \cdots \tag{4}$$

More precisely, we can define a Markov skeleton jump process $Z = \{Z(t), t \geq 0\}$ by the following procedure.

$$\tau_0 = 0\,, \quad \tau_n = \sum_{k=0}^{n-1} Y_k \quad \text{for } n \geq 1\,, \quad \zeta = \sum_{k \geq 0} Y_k\,. \tag{5}$$

$$Z(t) = X_n \quad \text{for } \tau_n \leq t < \tau_{n+1}\,, \quad \text{and} \quad Z(t) = \Delta \ \text{ for } t \geq \zeta\,. \tag{6}$$

In what follows we shall write $Z = (X, Y)$ if a Markov skeleton jump process Z is determined by (X, Y) with the above procedure. We denote by $\mathscr{F}^X$ the σ-algebra generated by $X = \{X_n, n \geq 0\}$. We are now in a position to introduce the notion of web Markov skeleton process, which is the main objective of this paper.

DEFINITION 2.3. Let $Z = (X, Y)$ be a Markov skeleton jump process described as above. Z is called a *web Markov skeleton process* (WMSP in short), if given the σ-algebra $\mathscr{F}^X$, the random variables Y_n $(n \geq 0)$ are conditionally independent to each other.

The notion of WMSP appeared recently in [9, 10]. The authors found that WMSP is a very suitable framework for modeling the user browsing behavior on the Web, and is a very useful mathematical tool for computing the web page importance. When modeling the user browsing behavior, the state space E is a collection of web pages, and the WMSP $Z = (X, Y)$ models a random walk on E. The random variable $Z(t)$ denotes the current page browsed by the user at time t. The process $X = \{X_n, n \geq 0\}$ describes the transition behavior between pages, which forms a Markov chain, and the sequence $Y = \{Y_n, n \geq 0\}$ represents staying times on the pages. The staying time on the current page is a random variable which may depend on the information of the current page and some other pages the user has visited or will visit, but in general will not depend on the other staying times. To our knowledge the framework of WMSP covers all the known stochastic processes used in studying the web page importance ranking. This is why we call it a web Markov skeleton process.

3. Scope of WMSP. The framework of WMSP is quite broad and convenient. It covers several existing classes of processes and contains also new classes of processes which are important either in applications or in theoretical study. Below we list a few special classes of WMSPs to outline an incomplete picture of its scope.

(A) Simple WMSP. We say that a WMSP $Z = (X, Y)$ is a simple web Markov skeleton process, if the conditional distribution of Y_n (given $\mathscr{F}^X$) depends only on X_n. More precisely, we propose the following definition.

DEFINITION 3.1. A WMSP $Z = (X, Y)$ is called a *simple web Markov skeleton process* (simple WMSP in short), if for all $t \geq 0$, it satisfies

$$P(Y_n \leq t \mid \mathscr{F}^X) = P(Y_n \leq t \mid X_n), \quad \text{for all } n \geq 0. \tag{7}$$

The class of simple WMSPs contains usual Markov processes as its special cases. See the two examples below for details.

- *Example* 1. Discrete-time Markov process. It is a simple WMSP $Z = (X, Y)$ such that Y_n is constant, i.e., $P(Y_n = 1) = 1$ for all n.
- *Example* 2. Continuous-time Markov process. Note that a continuous-time Markov process is always a Markov skeleton process in the sense of Definition 2.1 (e.g. let $\tau_n \equiv n$). Here we emphasize that if a continuous time Markv process Z is time homogeneous with right continuous sample paths on a discrete state space (i.e. Z is a Q-process), then Z is also a simple WMSP. In this case we may define τ_n to be the n-th jump of Z (set $\tau_0 = 0$) and let $X_n = Z_{\tau_n}$, $Y_n = \tau_{(n+1)} - \tau_n$. Conversely we have the following proposition.

PROPOSITION 3.2. *Let $Z = (X, Y)$ be a simple WMSP. Suppose that $X = \{X_n, n \geq 0\}$ is a time homogeneous Markov chain on a discrete space E, which satisfies $P(X_{n+1} \neq X_n) = 1$ for all $n \geq 0$, and that there exists a family of positive numbers $\{\lambda_x, x \in E\}$ such that*

$$P(Y_n \leq t \mid \mathscr{F}^X) = 1 - \exp\{-\lambda_{X_n} t\}, \quad \text{for all } n \geq 0.$$

Then Z is a time homogeneous continuous time Markov process.

We leave the proof of the above proposition to the readers.

We remark that the class of simple WMSP contains also processes which are not usual Markov processes. A simple WMSP is a special example of a semi-Markov process (cf. part (B) below). It is also a special example of a mirror semi-Markov process (cf. part (C) below). Moreover, in some circumstance we may reduce the study of a semi-Markov process or a mirror semi-Markov process to the study of a simple WMSP. We shall come back to this point later in Section 7.

(B) Semi-Markov process. In our framework, we shall call a WMSP $Z = (X, Y)$ a semi-Markov process, if the conditional distribution of Y_n depends only on X_n and X_{n+1}. More precisely, we propose the following definition.

DEFINITION 3.3. A WMSP $Z = (X, Y)$ is called a *semi-Markov process*, if for all $t \geq 0$, it satisfiies

$$P(Y_n \leq t \mid \mathscr{F}^X) = P(Y_n \leq t \mid X_n, X_{n+1}), \quad \text{for all } n \geq 0. \tag{8}$$

We shall show that our notion of semi-Markov processes is in fact more general than the one usually defined in the literature. To this end let us refer to the one defined in the literature as the *classical semi-Markov process*. A classical semi-Markov process Z is defined as a continuous time realization of a Markov renewal process. To be precise, let $X = \{X_n, n \geq 0\}$ be a sequence of random variables taking values in E, and $Y = \{Y_n, n \geq 0\}$ be a sequence of positive random variables. The sequence (X, Y) is called a *Markov renewal process*, if

$$P(X_{n+1} \in B, Y_n \leq t \mid X_0, \dots, X_n; Y_0, \dots, Y_{n-1}) = P(X_{n+1} \in B, Y_n \leq t \mid X_n), \tag{9}$$

for all $n \geq 0$, $B \in \mathscr{E}$, $t \geq 0$. A *classical semi-Markov process* $Z = \{Z(t), t \geq 0\}$ is then defined by the regime (5) and (6) with a Markov renewal process (X, Y). In the literature it is usually assumed further that the state space E is discrete and (X, Y) is time homogeneous. See e.g. [3, 6, 7, 25] and references therein for more details.

In the following theorem we do not require that (X, Y) is time homogeneous.

THEOREM 3.4. *Let Z be a classical semi-Markov process described as above. Then $Z = (X, Y)$ is a semi-Markov process in the sense of Definition* 3.3.

PROOF. By (9), it is clear that $X = \{X_n, n \geq 0\}$ is a Markov chain. We need only to check that given the σ-algebra $\mathscr{F}^X$, the random variables Y_n $(n \geq 0)$ are conditionally independent and their conditional distributions satisfy (8). The latter is a direct consequence of Lemma 3.5 below. □

The assertion (11) of Lemma 3.5 below might be known in the case where E is discrete and (X, Y) is time homogeneous. But we could not find a proof available to the stronger assertion (10) in the case where E is discrete and (X, Y) is not necessarily time homogeneous. For the sake of the rigor and also for the convenience of the readers, we include a complete proof here. We use $\boldsymbol{N}$ to denote the set of nonnegative integers.

LEMMA 3.5. *Suppose that (X, Y) is a Markov renewal process satisfying* (9) *with a discrete state space E. Denote by $\mathscr{F}_{\tau_n} = \sigma\{X_0, X_1, \dots, X_n; Y_0, Y_1, \dots, Y_{n-1}\}$ (by convention $\mathscr{F}_{\tau_0} := \sigma\{X_0\}$). Then for each $n \in \boldsymbol{N}$, $m \in \boldsymbol{N}$, $t_n \geq 0$, $t_{n+1} \geq 0, \dots, t_{n+m} \geq 0$, we have*

$$\begin{aligned} &P\big(Y_n \leq t_n, Y_{n+1} \leq t_{n+1}, \dots, Y_{n+m} \leq t_{n+m} \mid \mathscr{F}_{\tau_n} \vee \mathscr{F}^X\big) \\ &\quad = P\big(Y_n \leq t_n \mid X_n, X_{n+1}\big) P\big(Y_{n+1} \leq t_{n+1} \mid X_{n+1}, X_{n+2}\big) \times \cdots \\ &\qquad \times P\big(Y_{n+m} \leq t_{n+m} \mid X_{n+m}, X_{n+m+1}\big), \quad a.s. \end{aligned} \tag{10}$$

In particular, for each $m \in \boldsymbol{N}$, $t_0 \geq 0$, $t_1 \geq 0, \dots, t_m \geq 0$, we have

$$\begin{aligned} &P\big(Y_0 \leq t_0, Y_1 \leq t_1, \dots, Y_m \leq t_m \mid \mathscr{F}^X\big) \\ &\quad = P\big(Y_0 \leq t_0 \mid X_0, X_1\big) P\big(Y_1 \leq t_1 \mid X_1, X_2\big) \cdots P\big(Y_m \leq t_m \mid X_m, X_{m+1}\big), \quad a.s. \end{aligned} \tag{11}$$

PROOF. Denote by $\mathscr{B}(R^+)$ the Borel sets of $R^+ := [0, \infty)$. For $A_0 \in \mathscr{B}(R^+)$, $A_1 \in \mathscr{B}(R^+), \dots, A_n \in \mathscr{B}(R^+)$, $b_0 \in E, b_1 \in E, \dots, b_{n+1} \in E$, we have by (9) that

$$\begin{aligned}
&P\big(Y_n \in A_n, X_{n+1} = b_{n+1} \mid X_0 = b_0, \dots, X_n = b_n, Y_0 \in A_0, \dots, Y_{n-1} \in A_{n-1}\big) \\
&\quad = P\big(Y_n \in A_n, X_{n+1} = b_{n+1} \mid X_n = b_n\big) \\
&\quad = P\big(Y_n \in A_n \mid X_n = b_n, X_{n+1} = b_{n+1}\big) P\big(X_{n+1} = b_{n+1} \mid X_n = b_n\big).
\end{aligned}$$

For notational convenience, we denote by $F^0_{\tau_n} := \{X_0 = b_0, \dots, X_n = b_n, Y_0 \in A_0, \dots, Y_{n-1} \in A_{n-1}\}$ for the event in $\mathscr{F}_{\tau_n}$, and write the above formula in a concise form as follows,

$$\mathscr{P}\big(Y_n, X_{n+1} \mid F^0_{\tau_n}\big) = \mathscr{P}\big(Y_n, X_{n+1} \mid X_n\big) = \mathscr{P}\big(Y_n \mid X_n, X_{n+1}\big)\mathscr{P}\big(X_{n+1} \mid X_n\big),$$

without explicitly mentioning $A_0, A_1, \dots, A_n, b_0, b_1, \dots, b_{n+1}$. In the formulae (12), (13), (14) and (15) below we take the same convention without explicitly mentioning $A_k s$ and $b_k s$.

For any $q \in \boldsymbol{N}$ large enough, we have

$$\begin{aligned}
(12)\qquad &\mathscr{P}\big(Y_n, X_{n+k},\ 1 \le k \le q \mid F^0_{\tau_n}\big) \\
&\quad = \mathscr{P}\big(Y_n, X_{n+1} \mid F^0_{\tau_n}\big)\mathscr{P}\big(X_{n+k},\ 2 \le k \le q \mid F^0_{\tau_{(n+1)}}\big) \\
&\quad = \mathscr{P}\big(Y_n \mid X_n, X_{n+1}\big)\mathscr{P}\big(X_{n+1} \mid X_n\big)\mathscr{P}\big(X_{n+k},\ 2 \le k \le q \mid X_{n+1}\big) \\
&\quad = \mathscr{P}\big(Y_n \mid X_n, X_{n+1}\big)\mathscr{P}\big(X_{n+k},\ 1 \le k \le q \mid X_n\big).
\end{aligned}$$

Then for any $m \le q - 1$, we have

$$\begin{aligned}
(13)\quad &\mathscr{P}\big(Y_n, Y_{n+1}, \dots, Y_{n+m}, X_{n+k},\ 1 \le k \le q \mid F^0_{\tau_n}\big) \\
&\quad = \mathscr{P}\big(Y_n, X_{n+1} \mid F^0_{\tau_n}\big)\mathscr{P}\big(Y_{n+1}, \dots, Y_{n+m}, X_{n+k},\ 2 \le k \le q \mid F^0_{\tau_{(n+1)}}\big) \\
&\quad = \mathscr{P}\big(Y_n, X_{n+1} \mid F^0_{\tau_n}\big)\mathscr{P}\big(Y_{n+1}, X_{n+2} \mid F^0_{\tau_{(n+1)}}\big) \\
&\qquad \times \mathscr{P}\big(Y_{n+2}, \dots, Y_{n+m}, X_{n+k},\ 3 \le k \le q \mid F^0_{\tau_{(n+2)}}\big) \\
&\quad = \mathscr{P}\big(Y_n, X_{n+1} \mid F^0_{\tau_n}\big) \cdots \mathscr{P}\big(Y_{n+m-1}, X_{n+m} \mid F^0_{\tau_{(n+m-1)}}\big) \\
&\qquad \times \mathscr{P}\big(Y_{n+m}, X_{n+k},\ m+1 \le k \le q \mid F^0_{\tau_{(n+m)}}\big) \\
&\quad = \mathscr{P}\big(Y_n \mid X_n, X_{n+1}\big)\mathscr{P}\big(X_{n+1} \mid X_n\big) \cdots P\big(Y_{n+m-1} \mid X_{n+m-1}, X_{n+m}\big) \\
&\qquad \times \mathscr{P}\big(X_{n+m} \mid X_{n+m-1}\big)\mathscr{P}\big(Y_{n+m} \mid X_{n+m}, X_{n+m+1}\big) \\
&\qquad \times \mathscr{P}\big(X_{n+m+k},\ m+1 \le k \le q \mid X_{n+m}\big) \\
&\quad = \mathscr{P}\big(Y_n \mid X_n, X_{n+1}\big) \cdots \mathscr{P}\big(Y_{n+m} \mid X_{n+m}, X_{n+m+1}\big)\mathscr{P}\big(X_{n+k},\ 1 \le k \le q \mid X_n\big),
\end{aligned}$$

and on the other hand, we have

$$\begin{aligned}
(14)\quad &\mathscr{P}\big(Y_n, Y_{n+1}, \dots, Y_{n+m}, X_{n+k},\ 1 \le k \le q \mid F^0_{\tau_n}\big) \\
&\quad = \mathscr{P}\big(Y_n, Y_{n+1}, \dots, Y_{n+m} \mid F^0_{\tau_n}, X_{n+k},\ 1 \le k \le q\big)\mathscr{P}\big(X_{n+k},\ 1 \le k \le q \mid F^0_{\tau_n}\big) \\
&\quad = \mathscr{P}\big(Y_n, Y_{n+1}, \dots, Y_{n+m} \mid F^0_{\tau_n}, X_{n+k},\ 1 \le k \le q\big)\mathscr{P}\big(X_{n+k},\ 1 \le k \le q \mid X_n\big).
\end{aligned}$$

Comparing (13) and (14), we get

$$\begin{aligned}(15)\qquad &\mathscr{P}\big(Y_n, Y_{n+1}, \ldots, Y_{n+m} \mid F^0_{\tau_n}, X_{n+k},\ 1 \le k \le q\big)\\ &\quad = \mathscr{P}\big(Y_n \mid X_n, X_{n+1}\big) \cdots \mathscr{P}\big(Y_{n+m} \mid X_{n+m}, X_{n+m+1}\big)\,.\end{aligned}$$

In the original non-concise form, (15) means that for $m \le q-1$, $A_0 \in \mathscr{B}(R^+)$, $A_1 \in \mathscr{B}(R^+)$, $\ldots$, $A_{n+m} \in \mathscr{B}(R^+)$, $b_0 \in E$, $b_1 \in E, \ldots, b_{n+q} \in E$, we have

$$\begin{aligned}(16)\quad &P\big(Y_{n+l} \in A_{n+l},\ 0 \le l \le m \mid Y_h \in A_h,\ 0 \le h \le n-1; X_k = b_k,\ 0 \le k \le n+q\big)\\ &\quad = P\big(Y_n \in A_n \mid X_n = b_n, X_{n+1} = b_{n+1}\big) \times \cdots\\ &\qquad \times P\big(Y_{n+m} \in A_{n+m} \mid X_{n+m} = b_{n+m}, X_{n+m+1} = b_{n+m+1}\big)\,.\end{aligned}$$

Because $(R^+, \mathscr{B}(R^+))$ is separable, there exists $\{C_k\}_{k\ge 1} \subseteq \mathscr{B}(R^+)$ satisfying $\mathscr{B}(R^+) = \sigma(\{C_k\}_{k\ge1})$. For $s \in N$, $s \ge (m+1)$, we define $\mathscr{B}_s := \sigma(\{C_n\}_{1\le n\le s})$ and $\mathscr{G}_s := \sigma\big(Y_l^{-1}(\mathscr{B}_s)$, for all $0 \le l \le n-1$; $X_k^{-1}(\mathscr{E}),\ 0 \le k \le n+s\big)$. Then $\mathscr{G}_s$ contains at most countable many atoms, and each element in $\mathscr{G}_s$ is expressed as a union of some of the atoms. Therefore we can use (16) to get

$$\begin{aligned}(17)\qquad &P\big(Y_{n+l} \in A_{n+l}, 0 \le l \le m \big| \mathscr{G}_s\big)\\ &\quad = P\big(Y_n \in A_n \big| \sigma\big(X_n^{-1}(\mathscr{E}), X_{n+1}^{-1}(\mathscr{E})\big)\big) \times \cdots\\ &\qquad \times P\big(Y_{n+m} \in A_{n+m} \big| \sigma\big(X_{n+m}^{-1}(\mathscr{E}), X_{n+m+1}^{-1}(\mathscr{E})\big)\big) \quad a.s\,.\end{aligned}$$

Notice that $\{\mathscr{G}_s\}_{s\ge m+1}$ is an increasing family of sub-algebras and $\bigvee_{s\ge m+1} \mathscr{G}_s = \mathscr{F}_{\tau_n} \vee \mathscr{F}^X$. Letting $s \to \infty$ in (17), by the well known martingale convergence theorem (cf. e.g. [28]), we obtain the assertion (10). Letting $n = 0$ in (10) we get the assertion (11). □

(C) Mirror semi-Markov process. Mirror semi-Markov processes are a new class of processes in the framework of WMSP. We call a WMSP $Z = (X, Y)$ mirror semi-Markov process, if the conditional distribution of Y_n depends only on X_n and X_{n-1}. More precisely, we propose the following definition.

DEFINITION 3.6. A WMSP $Z = (X, Y)$ is called a *mirror semi-Markov process*, if for all $t \ge 0$, it satisfies

$$(18)\qquad P(Y_n \le t \mid \mathscr{F}^X) = \begin{cases} P(Y_n \le t \mid X_n, X_{n-1})\,, & \text{for all } n \ge 1\,,\\ P(Y_0 \le t \mid X_0)\,, & n = 0\,.\end{cases}$$

Mirror semi-Markov process was invented in [9, 10] and played an important role in modeling the browsing behavior on the mobile Web. It is a typical example of a Markov skeleton process in the sense of Definition 2.1 but not a Hou's Markov skeleton process (cf. Definition 2.2).

4. Time-homogeneous WMSP. Time-homogeneous web Markov skeleton processes are most important both in applications and in theoretical research. For concrete examples, such as semi-Markov process and mirror semi-Markov process, it is easy to describe time homogeneity. We found that it is not easy to do so for general WMSPs. In this section we shall introduce a general definition of time homogeneity for WMSP. The idea employed in our definition might be useful elsewhere.

To state our definition, we introduce first the notion of a *shift operator* on $\boldsymbol{N}$. For $m \in \boldsymbol{N}$, we define $\theta_m : \boldsymbol{N} \mapsto \boldsymbol{N}$ by setting $\theta_m(k) = m + k$. The operator θ_m is called the *m-step shift operator* on $\boldsymbol{N}$. For an arbitrary subset S of $\boldsymbol{N}$, we define $S \circ \theta_m$ by setting

$$S \circ \theta_m = \{\theta_m(k); k \in S\} = \{k + m; k \in S\}. \tag{19}$$

The operator $S \circ \theta_m$ is called the *m-step shift* of S. Further let S be an arbitrary finite subset of $\boldsymbol{N}$ with $\#\{k; k \in S\} = d$. We use $(X_k, k \in S)$ to denote a d-dimensional random vector consisting of random variables $\{X_k\}_{k \in S}$, with subscripts listed in ascending order. That is, $(X_k, k \in S) = (X_{k_1}, X_{k_2}, \dots, X_{k_d}) \in \tilde{E}^d$ with $k_i \in S$ for all i and $k_1 < k_2 < \cdots < k_d$. Here and henceforth $\tilde{E}^d := \bigotimes_{k=1}^d \tilde{E}_k$ with each $\tilde{E}_k$ identically equal to $\tilde{E}$. We shall use $\mathscr{B}(\tilde{E}^d)$ to denote the Borel sets of $\tilde{E}^d$.

Below is our definition of time homogeneity for WMSP.

DEFINITION 4.1. A WMSP $Z = (X, Y)$ is said to be time-homogeneous (after n_0), if its embedded Markov chain X is time-homogeneous, and for each $n \geq 0$ there exists a finite set $S_n \subset \boldsymbol{N}$ with $\max_{n \geq 0} \#\{k; k \in S_n\} < \infty$, satisfying the following properties:

$$P(Y_n \leq t | \mathscr{F}^X) = P\big(Y_n \leq t | (X_k, k \in S_n)\big), \tag{20}$$

$$S_{n+m} = S_n \circ \theta_m, \text{ for all } m \geq 0, \ n \geq n_0 := \max_{n \geq 0} \#\{k; k \leq n, \ k \in S_n\}, \tag{21}$$

$$\begin{gathered} P\big(Y_n \leq t | (X_k, k \in S_n) \in B\big) = P\big(Y_{n+m} \leq t | (X_k, k \in S_{n+m}) \in B\big), \\ \text{for all } \ m \geq 0, \ n \geq n_0, \ B \in \mathscr{B}(\tilde{E}^d), \quad (d := \#\{k; k \in S_n\}). \end{gathered} \tag{22}$$

In the study of time homogeneous WMSP, the concept of the *kernel* (cf. Remark 4.4 bellow) is quite useful. We are now going to describe it. Let Z be a time homogeneous WMSP and S_n be specified by the above definition. We set

$$S_n^+ := \{k > n; k \in S_n\}, \quad d_n^+ := \#\{k > n; k \in S_n\}, \tag{23}$$

and

$$S_n^- := \{k < n; k \in S_n\}, \quad d_n^- := \#\{k < n; k \in S_n\}. \tag{24}$$

By the property (21), when $n \geq n_0$, all the d_n^+'s are the same, which will be denoted by d^+. Also all the d_n^-'s for $n \geq n_0$ are the same, which will be denoted by d^-. For notational convenience, we shall make the convention that $\tilde{E}^0 = \emptyset$ and $\{(X_k, k \in \emptyset) \in A\} = \Omega$, where $\emptyset$ is the empty set. With the above convention, the expressions (25) and (26) below are still meaningful when $d^- = 0$.

Let $A \in \mathscr{B}(\tilde{E}^{d^-})$, $B \in \mathscr{B}(\tilde{E}^{d^+})$, $I \in \mathscr{B}(\tilde{E}) = \mathscr{E}$, and $t \geq 0$. We define for $n \geq n_0$,

$${}_AQ_{IB}(t) := P\big((X_l, l \in S_n^+) \in B, Y_n \leq t \mid X_n \in I, (X_k, k \in S_n^-) \in A\big), \tag{25}$$

$${}_AF_{IB}(t) := P\big(Y_n \leq t \mid X_n \in I, (X_k, k \in S_n^-) \in A, (X_l, l \in S_n^+) \in B\big). \tag{26}$$

PROPOSITION 4.2. *Both* ${}_AQ_{IB}(t)$ *and* ${}_AF_{IB}(t)$ *are independent of* n. *Moreover, if* $(n+1) \in S_n^+$, *then the probability distribution of* Z *is uniquely determined by the initial distribution of* $\{X_0, Y_0, X_1, \dots, Y_{n_0-1}, X_{n_0}\}$ *and the family of conditional probabilities*

$$\{{}_AQ_{IB}(t);\ A \in \mathscr{B}(\tilde{E}^{d^-}), B \in \mathscr{B}(\tilde{E}^{d^+}), I \in \mathscr{B}(\tilde{E}),\ t \geq 0\}. \tag{27}$$

PROOF. It follows directly from the property (22) that ${}_AF_{IB}(t)$ is independent of n. By (25) and the Markov property of X, we get

$$\begin{aligned} {}_AQ_{IB}(t) &= P\big(Y_n \leq t \mid X_n \in I, (X_k, k \in S_n^-) \in A, (X_l, l \in S_n^+) \in B\big) \\ &\quad \times P\big((X_l, l \in S_n^+) \in B \mid X_n \in I, (X_k, k \in S_n^-) \in A\big) \\ &= {}_AF_{IB}(t) \cdot P_{IB}, \end{aligned} \tag{28}$$

where $P_{IB} := P\big((X_l, l \in S_n^+) \in B \mid X_n \in I\big)$ is independent of n by the time homogeneity of X. Hence ${}_AQ_{IB}(t)$ is independent of n. Conversely, from (28) we get

$$P_{IB} = \lim_{t\to\infty} {}_AQ_{IB}(t) = P\big((X_l, l \in S_n^+) \in B \mid X_n \in I\big), \tag{29}$$

$${}_AF_{IB}(t) = \begin{cases} \frac{{}_AQ_{IB}(t)}{P_{IB}} & \text{if } P_{IB} > 0, \\ 1 & \text{if } P_{IB} = 0. \end{cases} \tag{30}$$

Therefore, $\{{}_AF_{IB}(t)\}$ and $\{P_{IB}\}$ are also uniquely determined by $\{{}_AQ_{IB}(t)\}$, from which the second assertion of the proposition follows. □

REMARK 4.3. If $(n+1) \in S_n^+$, from (29) one can re-capture the one step transition function of X by setting $B = B_1 \times \bigotimes_{k=2}^{d^+} \tilde{E}_k$ with B_1 running over all the Borel sets of $\tilde{E}$. In the case where $(n+1)$ is not in S_n^+, it is not clear whether or not we can re-capture the one step transition function of X. But if we set

$$\tilde{S}_n^+ := \big\{k > n; k \in S_n \cup \{(n+1)\}\big\}, \quad \tilde{d}^+ := \#\{k > n; k \in \tilde{S}_n^+\},$$

then we can re-capture the one step transition function of X from the family of conditional probabilities

$$\{{}_AQ_{IB}(t);\ A \in \mathscr{B}(\tilde{E}^{d^-}), B \in \mathscr{B}(\tilde{E}^{\tilde{d}^+}), I \in \mathscr{B}(\tilde{E}),\ t \geq 0\}. \tag{31}$$

Actually, all the assertions of Proposition 4.2 as well as the above Remark 4.3 remain true, if we replace S_n^+ by $\tilde{S}_n^+$, d^+ by $\tilde{d}^+$, and (27) by (31). Note that if $(n+1)$ is not in S_n^+, then ${}_AF_{IB}(t)$ will be independent of X_{n+1}. For example if $d^+ = 0$ and hence $\tilde{d}^+ = 1$, then (28) will become

$$\begin{aligned} {}_AQ_{IB}(t) &= P\big(X_{n+1} \in B, Y_n \leq t \mid X_n \in I, (X_k, k \in S_n^-) \in A\big) \\ &= P\big(Y_n \leq t \mid X_n \in I, X_{n+1} \in B, (X_k, k \in S_n^-) \in A\big) \\ &\quad \times P\big(X_{n+1} \in B \mid X_n \in I, (X_k, k \in S_n^+) \in A\big) \\ &= P\big(Y_n \leq t \mid X_n \in I, (X_l, l \in S_n^-) \in A\big) \times P\big(X_{n+1} \in B \mid X_n \in I\big) \\ &:= {}_AF_I(t) P_{IB}. \end{aligned}$$

REMARK 4.4. Because of the importance of the family of conditional probabilities (31), we shall call (31) the *kernel* of Z, which describes the statistical properties of $Z = (X, Y)$ after n_0. For a given kernel (31) as well as an initial distribution of X_0 and a family of conditional probability distributions $P(Y_n \le t \mid (X_k)_{k \in S_n})$ for $n = 0, 1, \ldots, n_0 - 1$, we can construct a unique (in weak sense) time homogeneous (after n_0) WMSP. The details will be discussed in our forthcoming paper [23].

EXAMPLE 4.5. Let $Z = (X, Y)$ be a time homogeneous WMSP with a discrete state space E. Denote the transition probabilities by

$$P_{ij} := P(X_{n+1} = j | X_n = i) = P(X_1 = j | X_0 = i)\,, \quad \text{for all } i, j \in \tilde{E}\,. \tag{32}$$

(A) If $S_n = \{n\}$, then Z is called a *time-homogeneous simple WMSP*. In this case if we define

$$F_i(t) = P(Y_n \le t \mid X_n = i)\,, \tag{33}$$

then the kernel of Z is expressed as

$$Q_{ij}(t) := P(X_{n+1} = j, Y_n \le t | \, X_n = i) = F_i(t) \cdot P_{ij}\,, \quad \text{for all } i, j \in \tilde{E}\,, \quad t \ge 0\,. \tag{34}$$

(B) If $S_n = \{n, n+1\}$, then Z is called a *time-homogeneous semi-Markov process*. In this case if we define

$$F_{ij}(t) = P(Y_n \le t \mid X_n = i, X_{n+1} = j)\,, \tag{35}$$

then the kernel of Z is expressed as

$$Q_{ij}(t) := P(X_{n+1} = j, Y_n \le t \mid X_n = i) = F_{ij}(t) \cdot P_{ij}\,, \quad \text{for all } i, j \in \tilde{E}\,, \quad t \ge 0\,. \tag{36}$$

(C) If $S_n = \{n-1, n\}$, then Z is called a *time-homogeneous mirror semi-Markov process*. In this case if we define

$${}_iF_j(t) = P(Y_n \le t \mid X_{n-1} = i, X_n = j)\,, \tag{37}$$

then the kernel of Z is expressed as

$$\begin{aligned} {}_iQ_{jk}(t) &:= P(X_{n+1} = k, Y_n \le t \mid X_{n-1} = i, X_n = j) \\ &= {}_iF_j(t) \cdot P_{jk}\,, \quad \text{for all } i, j, k \in \tilde{E}\,, \ t \ge 0\,. \end{aligned} \tag{38}$$

We shall discuss in more detail the class of time homogeneous mirror semi-Markov processes in the next section.

5. Mirror semi-Markov process. Throughout this section we assume that $Z = (X, Y)$ is a time homogeneous mirror semi-Markov process with a discrete state space E. For simplicity we assume further that $\zeta := \sum_{k \ge 0} Y_k = \infty$.

5.1. Further discussion on time homogeneity. We denote by $\boldsymbol{N}$ the set of nonnegative integers, and by $\boldsymbol{N}^+$ the set of positive integers.

LEMMA 5.1. *Let $m \in \boldsymbol{N}^+$, $s_1, \dots, s_m \in \boldsymbol{N}^+$, $1 \le s_1 < \cdots < s_m$ and $A_1, \dots, A_m \in \mathscr{B}(R^+)$. Suppose that $S \subset \boldsymbol{N}$ satisfies $\{0, \dots, s_1 - 1\} \cap S \neq \emptyset$. Then for any $\{i_k, k \in S\} \subset E$ and any $n \in \boldsymbol{N}^+$, we have*

$$\begin{aligned}&P\big[Y_{s_1} \in A_1, \dots, Y_{s_m} \in A_m \mid \{X_k = i_k\}_{k \in S}\big]\\&\quad = P\big[Y_{n+s_1} \in A_1, \dots, Y_{n+s_m} \in A_m \mid \{X_{n+k} = i_k\}_{k \in S}\big].\end{aligned}$$

PROOF. For notational convenience, we verify only a concrete case where $m = 2$, $s_1 = 2$, $s_2 = 4$, $S = \{1, 5\}$. The general case can be verified following the same idea. Thus we are going to verify

$$\begin{aligned}&P\big[Y_2 \in A_1, Y_4 \in A_2 \mid X_1 = i_1, X_5 = i_5\big]\\&\quad = P\big[Y_{n+2} \in A_1, Y_{n+4} \in A_2 \mid X_{n+1} = i_1, X_{n+5} = i_5\big].\end{aligned} \tag{39}$$

Denote by $C = \{X_{n+1} = i_1, X_{n+5} = i_5\}$. Without loss of generality we assume that $P(C) > 0$, otherwise the case will be negligible. By the definition of mirror semi-Markov processes we have

$$\begin{aligned}
&P\big[Y_{n+2} \in A_1, Y_{n+4} \in A_2 \mid X_{n+1} = i_1, X_{n+5} = i_5\big]\\
&\quad = \frac{1}{P(C)} \int_C P\big[Y_{n+2} \in A_1, Y_{n+4} \in A_2 \mid \mathscr{F}^X\big]\,\mathrm{d}P\\
&\quad = \frac{1}{P(C)} \int_C P\big[Y_{n+2} \in A_1 \mid \mathscr{F}^X\big] P\big[Y_{n+4} \in A_2 \mid \mathscr{F}^X\big]\,\mathrm{d}P\\
&\quad = \frac{1}{P(C)} \int_C P\big[Y_{n+2} \in A_1 \mid X_{n+1}, X_{n+2}\big] P\big[Y_{n+4} \in A_2 \mid X_{n+3}, X_{n+4}\big]\,\mathrm{d}P\\
&\quad = \frac{1}{P(C)} \int_C \Big(\sum_{i_2} P\big[Y_{n+2} \in A_1 \mid X_{n+1} = i_1, X_{n+2} = i_2\big] \times I_{\{X_{n+1} = i_1, X_{n+2} = i_2\}}\Big)\\
&\qquad\qquad \times \Big(\sum_{i_3, i_4} P\big[Y_{n+4} \in A_2 \mid X_{n+3} = i_3, X_{n+4} = i_4\big] \times I_{\{X_{n+3} = i_3, X_{n+4} = i_4\}}\Big)\mathrm{d}P\\
&\quad = \frac{1}{P(C)} \Big(\sum_{i_2, i_3, i_4} P\big[Y_{n+2} \in A_1 \mid X_{n+1} = i_1, X_{n+2} = i_2\big]\\
&\qquad\qquad \times P\big[Y_{n+4} \in A_2 \mid X_{n+3} = i_3, X_{n+4} = i_4\big]\\
&\qquad\qquad \times P\big[X_{n+1} = i_1, X_{n+2} = i_2, X_{n+3} = i_3, X_{n+4} = i_4, X_{n+5} = i_5\big]\Big)\\
&\quad = \sum_{i_2, i_3, i_4} P\big[Y_{n+2} \in A_1 \mid X_{n+1} = i_1, X_{n+2} = i_2\big]\\
&\qquad\qquad \times P\big[Y_{n+4} \in A_2 \mid X_{n+3} = i_3, X_{n+4} = i_4\big]
\end{aligned}$$

$$\times \frac{P\big[X_{n+2} = i_2, X_{n+3} = i_3, X_{n+4} = i_4, X_{n+5} = i_5 \mid X_{n+1} = i_1\big]}{P[X_{n+5} = i_5 \mid X_{n+1} = i_1]}$$

$$= \sum_{i_2,i_3,i_4} P\big[Y_2 \in A_1 \mid X_1 = i_1, X_2 = i_2\big] \times P\big[Y_4 \in A_2 \mid X_3 = i_3, X_4 = i_4\big]$$

$$\times \frac{P\big[X_2 = i_2, X_3 = i_3, X_4 = i_4, X_5 = i_5 \mid X_1 = i_1\big]}{P[X_5 = i_5 \mid X_1 = i_1]}.$$

In the above last step we used the time homogeneity assumption. By the same way we have

$$P\big[Y_2 \in A_1, Y_4 \in A_2 \mid X_1 = i_1, X_5 = i_5\big]$$

$$= \sum_{i_2,i_3,i_4} P\big[Y_2 \in A_1 \mid X_1 = i_1, X_2 = i_2\big] \times P\big[Y_4 \in A_2 \mid X_3 = i_3, X_4 = i_4\big]$$

$$\times \frac{P\big[X_2 = i_2, X_3 = i_3, X_4 = i_4, X_5 = i_5 \mid X_1 = i_1\big]}{P[X_5 = i_5 \mid X_1 = i_1]},$$

Hence (39) is verified. □

REMARK. The idea employed above is quite useful. Below we shall employ it again several times. The idea can also be employed for other classes of time homogeneous WMSPs, which will be discussed in our subsequent paper.

COROLLARY 5.2. *In the situation of Lemma* 5.1, *for any bounded Borel function f on $\boldsymbol{R}^m$, we have*

$$E\big[f(Y_{s_1}, Y_{s_2}, \ldots, Y_{s_m}) \mid \{X_k = i_k\}_{k\in S}\big]$$
$$= E\big[f(Y_{n+s_1}, Y_{n+s_2}, \ldots, Y_{n+s_m}) \mid \{X_{n+k} = i_k\}_{k\in S}\big].$$

PROOF. By the monotone class theorem. □

For $n \in \boldsymbol{N}^+$ and $A \in \mathscr{F}$, we define ${}_i\mathscr{P}_j^{(n)}(A) := P(A \mid X_{n-1} = i, X_n = j)$ provided that $P(X_{n-1} = i, X_n = j) > 0$. Then ${}_i\mathscr{P}_j^{(n)}$ is a probability measure on the space $(\Omega, \mathscr{F})$. For $t \geq 0$ we define $\tilde{Z}_t^{(n)} = Z_{\tau_n+t}$. We may regard $\tilde{Z}^{(n)} = \{\tilde{Z}_t^{(n)}, t \geq 0\}$ as a random variable taking values on $\mathfrak{R}_{\tilde{E}}[0, \infty)$, here and henceforth $\mathfrak{R}_{\tilde{E}}[0, \infty)$ denotes the sample space of all the $\tilde{E}$-valued right continuous and piecewise-constant functions.

THEOREM 5.3. *The random variable $\tilde{Z}^{(n)}$ on the probability space $(\Omega, \mathscr{F}, {}_i\mathscr{P}_j^{(n)})$ and the random variable $\tilde{Z}^{(1)}$ on the probability space $(\Omega, \mathscr{F}, {}_i\mathscr{P}_j^{(1)})$ induce the same probability distributions on the sample space $\mathfrak{R}_{\tilde{E}}[0, \infty)$.*

PROOF. For $m \in \boldsymbol{N}^+$, $t_1, \ldots, t_m \geq 0$ and $i_1, \ldots, i_m \in E$, we have

$${}_i\mathscr{P}_j^{(n)}(\tilde{Z}_{t_1}^{(n)} = i_1, \ldots, \tilde{Z}_{t_m}^{(n)} = i_m)$$
$$= P(Z_{(\tau_n+t_1)} = i_1, \ldots, Z_{(\tau_n+t_m)} = i_m \mid X_{n-1} = i, X_n = j)$$
$$= \sum_{l_1\in \boldsymbol{N},\ldots,l_m\in \boldsymbol{N}} P(Z_{(\tau_n+t_1)} = i_1, \ldots, Z_{(\tau_n+t_m)} = i_m, \tau_{(n+l_1)} - \tau_n \leq t_1 < \tau_{(n+l_1+1)} - \tau_n,$$
$$\ldots, \tau_{(n+l_m)} - \tau_n \leq t_m < \tau_{(n+l_m+1)} - \tau_n \mid X_{n-1} = i, X_n = j)$$

$$
\begin{aligned}
&= \sum_{l_1\in N,\dots,l_m\in N} P(X_{(n+l_1)} = i_1, \dots, X_{(n+l_m)} = i_m, \tau_{(n+l_1)} - \tau_n \le t_1 < \tau_{(n+l_1+1)} - \tau_n,\\
&\qquad \dots, \tau_{(n+l_m)} - \tau_n \le t_m < \tau_{(n+l_m+1)} - \tau_n \mid X_{n-1} = i, X_n = j)\\
&= \sum_{l_1\in N,\dots,l_m\in N} P(\tau_{(n+l_1)} - \tau_n \le t_1 < \tau_{(n+l_1+1)} - \tau_n, \dots, \tau_{(n+l_m)} - \tau_n \le t_m < \tau_{(n+l_m+1)} - \tau_n\\
&\qquad \mid X_{n-1} = i, X_n = j, X_{(n+l_1)} = i_1, \dots, X_{(n+l_m)} = i_m)\\
&\qquad \times P(X_{(n+l_1)} = i_1, \dots, X_{(n+l_m)} = i_m \mid X_{n-1} = i, X_n = j)\\
&= \sum_{l_1\in N,\dots,l_m\in N} P(\tau_{(1+l_1)} - \tau_1 \le t_1 < \tau_{(2+l_1)} - \tau_1, \dots, \tau_{(1+l_m)} - \tau_1 \le t_m < \tau_{(2+l_m)} - \tau_1\\
&\qquad \mid X_0 = i, X_1 = j, X_{(1+l_1)} = i_1, \dots, X_{(1+l_m)} = i_m)\\
&\qquad \times P(X_{(1+l_1)} = i_1, \dots, X_{(1+l_m)} = i_m \mid X_0 = i, X_1 = j)\\
&= \sum_{l_1\in N,\dots,l_m\in N} P(X_{(1+l_1)} = i_1, \dots, X_{(1+l_m)} = i_m, \tau_{(1+l_1)} - \tau_1 \le t_1 < \tau_{(1+l_1+1)} - \tau_1,\\
&\qquad \dots, \tau_{(1+l_m)} - \tau_1 \le t_m < \tau_{(1+l_m+1)} - \tau_1 \mid X_0 = i, X_1 = j)\\
&= \sum_{l_1\in N,\dots,l_m\in N} P(Z_{(\tau_1+t_1)} = i_1, \dots, Z_{(\tau_1+t_m)} = i_m, \tau_{(1+l_1)} - \tau_1 \le t_1 < \tau_{(1+l_1+1)} - \tau_1,\\
&\qquad \dots, \tau_{(1+l_m)} - \tau_1 \le t_m < \tau_{(1+l_m+1)} - \tau_1 \mid X_0 = i, X_1 = j)\\
&= P(Z_{(\tau_1+t_1)} = i_1, \dots, Z_{(\tau_1+t_m)} = i_m \mid X_0 = i, X_1 = j)\\
&= {}_i\mathscr{P}_j^{(1)}(\tilde{Z}_{t_1}^{(1)} = i_1, \dots, \tilde{Z}_{t_m}^{(1)} = i_m).
\end{aligned}
$$

In the middle step of the above deduction we employed Corollary 5.2. The proof of the theorem is completed by applying the monotone class theorem. □

5.2. Renewal theory. Thanks to the results obtained in the above subsection, we can now develop a renewal theory for mirror semi-Markov processes which is of importance in various applications.

In what follows we define

$$D^{(n)}(k) := \inf\{t \ge Y_n; \tilde{Z}_t^{(n)} = k\}, \tag{40}$$

and we call $D^{(n)}(k)$ the *first entry time* of $\tilde{Z}^{(n)}$ into the state k.

PROPOSITION 5.4. *For $i, j, k \in E$, $t \ge 0$ and $m, n \in N^+$, we have*

$${}_i\mathscr{P}_j^{(1)}(D^{(1)}(k) \le t) = {}_i\mathscr{P}_j^{(n)}(D^{(n)}(k) \le t) := {}_iG_{jk}(t), \tag{41}$$

$${}_i\mathscr{P}_j^{(1)}\left(X_{m+1} = k, \sum_{s=1}^{m} Y_s \le t\right) = {}_i\mathscr{P}_j^{(n)}\left(X_{m+n} = k, \sum_{s=n}^{m+n-1} Y_s \le t\right) := {}_iQ_{jk}^{(m)}(t), \tag{42}$$

$${}_i\mathscr{P}_j^{(1)}(\tilde{Z}_t^{(1)} = k) = {}_i\mathscr{P}_j^{(n)}(\tilde{Z}_t^{(n)} = k) := {}_if_{jk}(t). \tag{43}$$

PROOF. All the assertions follow from Theorem 5.3 and Corollary 5.2. We omit the details. □

Note that ${}_iQ^{(1)}_{jk}(t) := {}_iQ_{jk}(t)$ coincides with the kernel of Z expressed by (38).

For the sake of future references, we rewrite the formula (37) with our present notations in the following form.

$$ {}_i\mathscr{P}^{(1)}_j(Y_1 \le t) = {}_i\mathscr{P}^{(n)}_j(Y_n \le t) := {}_iF_j(t)\,. \tag{44} $$

The following lemma will be used in developing our renewal equations and renewal functional.

LEMMA 5.5. *For $A_1, A_2 \in \mathscr{B}(R^+)$ and $t \ge 0$, we have*

$$ \begin{aligned} {}_i\mathscr{P}^{(1)}_j\Bigg(X_{n+2} = k, \sum_{m=2}^{n+1} Y_m \in A_2 \mid X_2 = l, Y_1 \in A_1\Bigg) \\ = {}_j\mathscr{P}^{(2)}_l\Bigg(X_{n+2} = k, \sum_{m=2}^{n+1} Y_m \in A_2\Bigg), \end{aligned} \tag{45} $$

$$ {}_i\mathscr{P}^{(1)}_j(\tilde{Z}^{(2)}_t = k \mid X_2 = l, Y_1 \in A_1) = {}_j\mathscr{P}^{(2)}_l(\tilde{Z}^{(2)}_t = k), \tag{46} $$

$$ {}_i\mathscr{P}^{(1)}_j(D^{(2)}(k) \le t \mid X_2 = l, Y_1 \in A_1) = {}_j\mathscr{P}^{(2)}_l(D^{(2)}(k) \le t)\,, \quad \textit{for all } l \ne k\,. \tag{47} $$

PROOF. We prove only (45). The other two assertions can be proved similarly. We have

$$ \begin{aligned} &{}_i\mathscr{P}^{(1)}_j\Bigg(X_{n+2} = k, \sum_{m=2}^{n+1} Y_m \in A_2 \mid X_2 = l, Y_1 \in A_1\Bigg) \\ &= P\Bigg(X_{n+2} = k, \sum_{m=2}^{n+1} Y_m \in A_2 \mid X_0 = i, X_1 = j, X_2 = l, Y_1 \in A_1\Bigg) \\ &= P\Bigg(Y_1 \in A_1, \sum_{m=2}^{n+1} Y_m \in A_2 \mid X_0 = i, X_1 = j, X_2 = l, X_{n+2} = k\Bigg) \\ &\quad \times \frac{P(X_0 = i, X_1 = j, X_2 = l, X_{n+2} = k)}{P(X_0 = i, X_1 = j, X_2 = l, Y_1 \in A_1)} \\ &= P(Y_1 \in A_1 \mid X_0 = i, X_1 = j) \times P\Bigg(\sum_{m=2}^{n+1} Y_m \in A_2 \mid X_1 = j, X_2 = l, X_{n+2} = k\Bigg) \\ &\quad \times \frac{P(X_{n+2} = k \mid X_1 = j, X_2 = l)}{P(Y_1 \in A_1 \mid X_0 = i, X_1 = j)} \\ &= P(X_{n+2} = k \mid X_1 = j, X_2 = l) \times P\Bigg(\sum_{m=2}^{n+1} Y_m \in A_2 \mid X_1 = j, X_2 = l, X_{n+2} = k\Bigg) \\ &= P\Bigg(X_{n+2} = k, \sum_{m=2}^{n+1} Y_m \in A_2 \mid X_1 = j, X_2 = l\Bigg) \end{aligned} $$

$$= {}_j\mathscr{P}_l^{(2)}\left(X_{n+2} = k, \sum_{m=2}^{n+1} Y_m \in A_2\right).$$

Thus assertion (45) is proved. □

In what follows we shall frequently use the following notations. For $t > 0$, we take a partition $\{u_h\}_{0\le h\le s}$ of the interval $[0, t]$, i.e., $0 = u_0 < u_1 < \cdots < u_s < u_{(s+1)} = t$. Write $\Delta u_h := (u_h, u_{(h+1)}]$, for all $0 \le h \le s$, $\delta u := \max_h |u_{(h+1)} - u_h|$, and $\Delta[{}_iQ_{jk}(u_h)] := {}_iQ_{jk}(u_{(h+1)}) - {}_iQ_{jk}(u_h)$. Write $\delta_{jk} = 1$ if $j = k$, and $\delta_{jk} = 0$ if $j \ne k$.

Below is our renewal equations concerning ${}_if_{jk}(t)$ and ${}_iG_{jk}(t)$.

THEOREM 5.6. (i) ${}_if_{jk}(t) = \delta_{jk}[1 - {}_iF_j(t)] + \sum_{l\in E}\int_0^t[{}_jf_{lk}(t-u)]\mathrm{d}[{}_iQ_{jl}(u)]$;
(ii) ${}_iG_{jk}(t) = {}_iQ_{jk}(t) + \sum_{l\ne k}\int_0^t[{}_jG_{lk}(t-u)]\mathrm{d}[{}_iQ_{jl}(u)]$.

PROOF. For the assertion (i), employing the right continuity of $\{\tilde{Z}_t^{(1)}, t \ge 0\}$, and applying (46), we get

$$\begin{aligned}
{}_if_{jk}(t) &= {}_i\mathscr{P}_j^{(1)}(\tilde{Z}_t^{(1)} = k)\\
&= {}_i\mathscr{P}_j^{(1)}(\tilde{Z}_t^{(1)} = k, Y_1 > t) + {}_i\mathscr{P}_j^{(1)}(\tilde{Z}_t^{(1)} = k, Y_1 \le t)\\
&= \delta_{jk}[1 - {}_iF_j(t)] + \sum_{l\in E}\lim_{\delta u\to 0}\sum_{h=0}^{s}{}_i\mathscr{P}_j^{(1)}(\tilde{Z}^{(1)}_{(Y_1+t-u_h)} = k, Y_1 \in \Delta u_h, X_2 = l)\\
&= \delta_{jk}[1 - {}_iF_j(t)] + \sum_{l\in E}\lim_{\delta u\to 0}\sum_{h=0}^{s}\big({}_j\mathscr{P}_l^{(2)}(\tilde{Z}^{(2)}_{(t-u_h)} = k)\Delta[{}_iQ_{jl}(u_h)]\big)\\
&= \delta_{jk}[1 - {}_iF_j(t)] + \sum_{l\in E}\int_0^t[{}_jf_{lk}(t-u)]\mathrm{d}[{}_iQ_{jl}(u)].
\end{aligned}$$

For the assertion (ii), we can employ (47) to get

$$\begin{aligned}
{}_iG_{jk}(t) &= {}_i\mathscr{P}_j^{(1)}(D^{(1)}(k) \le t)\\
&= {}_i\mathscr{P}_j^{(1)}(D^{(1)}(k) \le t, X_2 = k) + \sum_{l\ne k}{}_i\mathscr{P}_j^{(1)}(D^{(1)}(k) \le t, X_2 = l)\\
&= {}_i\mathscr{P}_j^{(1)}(Y_1 \le t, X_2 = k)\\
&\quad + \sum_{l\ne k}\lim_{\delta u\to 0}\sum_{h=0}^{s}{}_i\mathscr{P}_j^{(1)}(D^{(1)}(k) - Y_1 \le t - u_h, X_2 = l, Y_1 \in \Delta u_h)\\
&= {}_iQ_{jk}(t) + \sum_{l\ne k}\lim_{\delta u\to 0}\sum_{h=0}^{s}\big({}_j\mathscr{P}_l^{(2)}(D^{(2)}(k) \le t - u_h)\Delta[{}_iQ_{jl}(u_h)]\big)\\
&= {}_iQ_{jk}(t) + \sum_{l\ne k}\int_0^t[{}_jG_{lk}(t-u)]\mathrm{d}[{}_iQ_{jl}(u)].
\end{aligned}$$

□

Denote by $\tilde{N}_t(k)$ the number of visits to the state k by the process $\tilde{Z}^{(1)}$ during the interval $(0, t]$. We define the *renewal functional* $M(i, j, k; t)$ of Z as the expectation of $\tilde{N}_t(k)$ with respect to the probability ${}_i\mathscr{P}_j^{(1)}$, namely,

$$M(i, j, k; t) := E[\tilde{N}_t(k) \mid X_0 = i, X_1 = j].$$

Below are our results about the renewal functional.

THEOREM 5.7. (i) $M(i, j, k; t) = \sum_{n=1}^{\infty} {}_iQ_{jk}^{(n)}(t)$.

(ii) ${}_iQ_{jk}^{(n+1)}(t) = \sum_{l\in E} \int_0^t [{}_jQ_{lk}^{(n)}(t-u)]\mathrm{d}[{}_iQ_{jl}(u)]$.

PROOF. To prove the assertion (i), we write $A_n = I_{\{X_{n+1}=k,\ \sum_{m=1}^n Y_m \le t\}}$ for $n \ge 1$. Then $\tilde{N}_t(k) = \sum_{n=1}^{\infty} A_n$. Therefore,

$$\begin{aligned} M(i, j, k; t) &= E[\tilde{N}_t(k) \mid X_0 = i, X_1 = j] \\ &= \sum_{n=1}^{\infty} E[A_n \mid X_0 = i, X_1 = j] \\ &= \sum_{n=1}^{\infty} {}_i\mathscr{P}_j^{(1)}\left(X_{n+1} = k, \sum_{m=1}^{n} Y_m \le t\right) \\ &= \sum_{n=1}^{\infty} {}_iQ_{jk}^{(n)}(t)\,. \end{aligned}$$

To prove the recursion equation (ii), we have

$$\begin{aligned} {}_iQ_{jk}^{(n+1)}(t) &= {}_i\mathscr{P}_j^{(1)}\left(X_{n+2} = k, \sum_{m=1}^{n+1} Y_m \le t\right) \\ &= \sum_{l\in E} {}_i\mathscr{P}_j^{(1)}\left(X_2 = l, X_{n+2} = k, \sum_{m=1}^{n+1} Y_m \le t\right). \end{aligned} \tag{48}$$

Applying (45), we get

$$\begin{aligned} {}_iQ_{jk}^{(n+1)}(t) &= \sum_{l\in E} \lim_{\delta u\to 0} \sum_{h=0}^{s} {}_i\mathscr{P}_j^{(1)}\left(X_2 = l, X_{n+2} = k, Y_1 \in \Delta u_h, \sum_{m=2}^{n+1} Y_m \le t - u_h\right) \\ &= \sum_{l\in E} \lim_{\delta u\to 0} \sum_{h=0}^{s} \left({}_i\mathscr{P}_j^{(1)}\left(X_{n+2} = k, \sum_{m=2}^{n+1} Y_m \le t - u_h \mid X_2 = l, Y_1 \in \Delta u_h\right)\right. \\ &\qquad\qquad \left.\times\ {}_i\mathscr{P}_j^{(1)}(X_2 = l, Y_1 \in \Delta u_h)\right) \\ &= \sum_{l\in E} \lim_{\delta u\to 0} \sum_{h=0}^{s} \left({}_j\mathscr{P}_l^{(2)}\left(X_{n+2} = k, \sum_{m=2}^{n+1} Y_m \le t - u_h\right) {}_i\mathscr{P}_j^{(1)}(X_2 = l, Y_1 \in \Delta u_h)\right) \end{aligned}$$

$$= \sum_{l \in E} \lim_{\delta u \to 0} \sum_{h=0}^{s} \left({}_j Q_{lk}^{(n)}(t - u_h) \Delta[{}_i Q_{jl}(u_h)] \right)$$

$$= \sum_{l \in E} \int_0^t [{}_j Q_{lk}^{(n)}(t - u)] \mathrm{d}[{}_i Q_{jl}(u)] . \qquad \square$$

5.3. Staying times and first entry times. We write ${}_i T_j$ and ${}_i V_{jk}$ for Y_1 and $D^{(1)}(k)$ respectively when they are viewed as the random variables on the probability space $(\Omega, \mathscr{F}, {}_i \mathscr{P}_j^{(1)})$. We call ${}_i T_j$ the *staying time* on the state j with the knowledge of the previous state i, and call ${}_i V_{jk}$ the *first entry time* into the state k from the present state j with the knowledge of the previous state i. By (44) and (41) we know that their distributions are ${}_i F_j(t)$ and ${}_i G_{jk}(t)$. We denote their expectations by ${}_i \alpha_j$ and ${}_i \mu_{jk}$ respectively. With the renewal equations obtained in the last subsection we can establish the following relation.

THEOREM 5.8. ${}_i \mu_{jk} = {}_i \alpha_j + \sum_{l \neq k} P_{jl} \cdot {}_j \mu_{lk}$.

PROOF. Taking the Laplace transforms of both sides of the assertion (ii) in Theorem 5.6, we get

$${}_i G_{jk}^*(s) = {}_i Q_{jk}^*(s) + \sum_{l \neq k} {}_j G_{lk}^*(s) \cdot {}_i Q_{jl}^*(s) .$$

By (38), we have ${}_i Q_{jk}(t) = {}_i F_j(t) \cdot P_{jk}$. Therefore

$${}_i G_{jk}^*(s) = {}_i F_j^*(s) \cdot P_{jk} + \sum_{l \neq k} {}_j G_{lk}^*(s) \cdot {}_i F_j^*(s) \cdot P_{jl} .$$

Taking the derivative for s and then letting $s = 0$, we obtain

$$\begin{aligned} {}_i \mu_{jk} &= P_{jk} \cdot {}_i \alpha_j + \sum_{l \neq k} ({}_j \mu_{lk} + {}_i \alpha_j) P_{jl} \\ &= \sum_{l \in E} P_{jl} \cdot {}_i \alpha_j + \sum_{l \neq k} {}_j \mu_{lk} \cdot P_{jl} \\ &= {}_i \alpha_j + \sum_{l \neq k} P_{jl} \cdot {}_j \mu_{lk} . \end{aligned} \qquad \square$$

Sometimes we need to study the staying times and the first entry times without knowing the previous states. But for a mirror semi-Markov process the conditional distribution of Y_n in general depends on both the current state X_n and the previous state X_{n-1}, which brings us some extra difficulty. To deal with it we introduce the following quantity:

$$c_{ij}^{(n)} := P(X_{n-1} = i \mid X_n = j) , \quad \text{for all } n \geq 1 , \ i \in E , \ j \in E , \tag{49}$$

provided that $P(X_n = j) > 0$. Note that if $P(X_n = j) > 0$, then $\{c_{ij}^{(n)}\}_{i \in E}$ forms a probability distribution on E. We call $c_{ij}^{(n)}$ the *contribution probability* from the state i to the state j (at the n step). The concept of contribution probability was first introduced and employed in [9].

In practice we often wish the contribution probabilities are independent of n. In this aspect the following simple lemma is useful.

LEMMA 5.9. *Suppose that the embedded Markov chain X admits a stationary distribution $\Pi = (\pi_i)_{i\in E}$ with $\pi_i > 0$ for all $i \in E$, and the initial distribution of X_0 is Π. Then for each $i, j \in E$, $c_{ij}^{(n)} =: c_{ij}$ is independent of n.*

PROOF. Let P_{ij} be the transition probability of X. Then by the assumption of the lemma we have

$$c_{ij} := c_{ij}^{(n)} = P(X_{n-1} = i \mid X_n = j) = \frac{P_{ij}P(X_{n-1} = i)}{P(X_n = j)} = \frac{P_{ij}\pi_i}{\pi_j}, \tag{50}$$

which is independent of n. □

When the contribution probabilities $c_{ij}^{(n)}$s are independent of n, the staying times and the first entry times without knowing the previous states will be also independent of n. We summarize it in the proposition below.

PROPOSITION 5.10. *Suppose that for each $i, j \in E$, $c_{ij}^{(n)} =: c_{ij}$ is independent of n. Then the following assertions hold.*

$$F_j(t) := P(Y_1 \le t \mid X_1 = j) = P(Y_n \le t \mid X_n = j) = \sum_{i\in E} {}_iF_j(t) \cdot c_{ij}\,. \tag{51}$$

$$G_{jk}(t) := P(D^{(1)}(k) \le t \mid X_1 = j) = P(D^{(n)}(k) \le t \mid X_n = j) = \sum_{i\in E} {}_iG_{jk}(t) \cdot c_{ij}\,. \tag{52}$$

PROOF. We prove only the first assertion, the second assertion can be proved similarly.

$$\begin{aligned} P(Y_n \le t \mid X_n = j) &= \sum_{i\in E} P(Y_n \le t, X_{n-1} = i \mid X_n = j) \\ &= \sum_{i\in E} P(Y_n \le t \mid X_{n-1} = i, X_n = j)P(X_{n-1} = i \mid X_n = j) \\ &= \sum_{i\in E} P(Y_1 \le t \mid X_0 = i, X_1 = j)P(X_0 = i \mid X_1 = j) \\ &= \sum_{i\in E} {}_iF_j(t) \cdot c_{ij}\,. \end{aligned}$$ □

In what follows we assume that the contribution probabilities $c_{ij}^{(n)}$s are independent of n. We write T_j for Y_1 when it is viewed as a random variable under the probability $P(\cdot \mid X_1 = j)$, and write V_{jk} for $D^{(1)}(k)$ when it is under the probability $P(\cdot \mid X_1 = j)$. We call T_j the staying time on the state j without knowing the previous state, and call V_{jk} the first entry time into the state k from the present state j without knowing the previous state. We denote their distributions by $F_j(t)$ and $G_{jk}(t)$ (cf. the above proposition), and denote their expectations by α_j and μ_{jk} respectively.

The results below will be used in the next section.

THEOREM 5.11. *We have the following relations among the expectations.*

(i) $\mu_{jk} = \sum_{i\in E} c_{ij} \cdot {}_i\mu_{jk}$,
(ii) $\alpha_j = \sum_{i\in E} c_{ij} \cdot {}_i\alpha_j$,
(iii) $\mu_{jk} = \alpha_j + \sum_{l\neq k} P_{jl} \cdot {}_j\mu_{lk}$.

If in addition the assumption of Lemma 5.9 *is fulfilled*, *then we have further*

(iv) $\mu_{jj} = \frac{\sum_{i\in E}\alpha_i\pi_i}{\pi_j}$.

PROOF. The assertions (i) and (ii) are direct consequences of (51) and (52), and the assertion (iii) follows from (i), (ii) and Theorem 5.8. By the assertion (iii) and the formula (50) we have,

$$\begin{aligned}\sum_{j\in E}\pi_j \cdot \mu_{jk} &= \sum_{j\in E}\pi_j \cdot \alpha_j + \sum_{j\in E}\sum_{l\neq k}\pi_j \cdot P_{jl} \cdot {}_j\mu_{lk}\\ &= \sum_{j\in E}\pi_j \cdot \alpha_j + \sum_{j\in E}\sum_{l\neq k}\pi_l \cdot c_{jl} \cdot {}_j\mu_{lk}\\ &= \sum_{j\in E}\pi_j \cdot \alpha_j + \sum_{l\neq k}\pi_l \cdot \mu_{lk}\,.\end{aligned}$$

Namely $\sum_{j\in E}\pi_j \cdot \alpha_j = \pi_k \cdot \mu_{kk}$, thus the assertion (iv) is obtained. □

6. Two results about limit distributions. In this section we discuss two results about the limit distributions of WMSPs. We show that under certain conditions the limit distributions of WMSPs are equal to the proportion of their average sojourn times. Hence the limit distributions play essential roles in computing page importance on the Web.

6.1. Limit distributions for semi-Markov processes. In this subsection we assume that $Z = (X, Y)$ is a time homogeneous semi-Markov process with a discrete state space E. Assume further that $\zeta := \sum_{k\geq 0} Y_k = \infty$ and X is an irreducible, recurrent Markov chain. For fixed state $j \in E$, we define

$$\begin{aligned}\xi_0 &= \inf\{n \geq 0; X_n = j\}\,,\\ \xi_1 &= \inf\{n > \xi_0; X_n = j\}\,,\\ &\dots\\ \xi_n &= \inf\{n > \xi_{n-1}; X_n = j\}\,,\\ &\dots .\end{aligned}$$

Recall that $\tau_n = \sum_{k=0}^{n-1} Y_k$.

LEMMA 6.1. *Define* $V_{jj}^{(n)} = \tau_{\xi_{(n+1)}} - \tau_{\xi_n}$ *and* $T_j^{(n)} = \tau_{(\xi_n+1)} - \tau_{\xi_n} = Y_{\xi_n}$. *Then* $\{V_{jj}^{(n)}, n \geq 0\}$ *are i.i.d. random variables*, *and* $\{T_j^{(n)}, n \geq 0\}$ *are i.i.d. random variables.*

PROOF. The assertions are more or less classical. We give only a sketch of the proof. It is easy to check that $\{\xi_n, n \geq 0\}$ is a set of stopping times of X. By the strong Markov property and time homogeneity of X we know that $\xi_1 - \xi_0, \xi_2 - \xi_1, \dots$ are i.i.d. random variables. Note that $\xi_1, \xi_2, \dots$ are all $\mathscr{F}^X$ measurable and $Y_1, Y_2, \dots$ are conditionally independent to each other given $\mathscr{F}^X$, from which together with the time homogeneity of Z we can obtain the desired assertions. $\square$

We denote by $F_j(t)$ and $G_{jj}(t)$ the distributions of $T_j^{(1)}$ and $V_{jj}^{(1)}$, and by α_j and μ_{jj} their expectations respectively. The readers may refer to the last subsection, in particular (51) and (52), to compare the corresponding concepts and notations for mirror semi-Markov processes.

Theorem 6.2 below is also more or less classical (cf. e.g. p. 217–p. 219 in [7]). We include a complete proof here because it is our base for deducing the corresponding result on mirror semi-Markov processes in the next subsection.

THEOREM 6.2. *For fixed $j \in E$, if $V_{jj}^{(1)}$ is not a lattice random variable and $\mu_{jj} = EV_{jj}^{(1)} < \infty$, then $f_j := \lim_{t\to\infty} P(Z(t) = j)$ exists and*

$$f_j = \frac{\alpha_j}{\mu_{jj}} = \lim_{t\to\infty} \frac{\int_0^t 1_{\{Z_s=j\}}\mathrm{d}s}{t}, \quad a.s. \tag{53}$$

If in addition X admits a unique limit distribution $\Pi = (\pi_i)_{i\in E}$, then we have further

$$f_j = \frac{\alpha_j \pi_j}{\sum_{i\in E} \alpha_i \pi_i}. \tag{54}$$

PROOF. Using the notation $\tilde{Z}_t = Z_{\tau_{\xi_0}+t}$, we can regard the process $\{\tilde{Z}_t, t \geq 0\}$ as a renewal process which starts from the state j and returns to the state j when the next renewal occurs. The lengths of time period between two successive renewals form a family of i.i.d random variables $\{V_{jj}^{(n)}\}_{n\geq 0}$, and the variable $V_{jj}^{(n)}$ can be divided into two successive parts: $V_{jj}^{(n)} = T_j^{(n)} + T_{\neq j}^{(n)}$, where $T_{\neq j}^{(n)}$ denotes the length of the time period within the n^{th} circle during which the process $\tilde{Z}_t$ is not on the state j. In this way, $\{\tilde{Z}_t, t \geq 0\}$ can be regarded as an alternating renewal process taking two values 'j ' and 'non j' alternatively in the periods $T_j^{(0)}, T_{\neq j}^{(0)}, T_j^{(1)}, T_{\neq j}^{(1)}, \ldots$. By the theory of alternating renewal processes (cf. e.g. p. 169–p. 170 in [7] and p. 286–p. 290 in [25]), we have

$$\lim_{t\to\infty} P(Z_t = j) = \lim_{t\to\infty} P(\tilde{Z}_t = j) = \frac{E(T_j^{(1)})}{E(T_j^{(1)}) + E(T_{\neq j}^{(1)})} = \frac{E(T_j^{(1)})}{E(V_{jj}^{(1)})} = \frac{\alpha_j}{\mu_{jj}}.$$

Denote by $N_t^j = \#\{n \in \boldsymbol{N}; X_n = j, \tau_n < t\}$. Applying the strong law of large numbers we get

$$\frac{\int_0^t 1_{\{Z_s=j\}}ds}{t} = \frac{\int_0^t 1_{\{Z_s=j\}}\mathrm{d}s}{N_t^j}\frac{N_t^j}{t} \xrightarrow{t\to\infty} \frac{\alpha_j}{\mu_{jj}} \ a.s.,$$

verifying the assertion (53).

We now verify the assertion (54). Denote by $N_j^{(n)}$ the number of times that X visits the state j in the past n transitions. By the ergodicity theory we know $\pi_j = \lim_{n\to\infty} N_j^{(n)}/n \ a.s.$ Let $p_j^{(n)}$ be the proportion of staying time on the state j during the previous n transitions.

According to (53), we have $f_j = \lim_{n\to\infty} p_j^{(n)}$ $a.s.$ Recall that $T_j^{(k)}$ represents the staying time of the k-th visit on the state j, we may rewrite $p_j^{(n)}$ as

$$p_j^{(n)} = \frac{\sum_{k=1}^{N_j^{(n)}} T_j^{(k)}}{\sum_i \sum_{k=1}^{N_i^{(n)}} T_i^{(k)}} = \frac{\frac{N_j^{(n)}}{n} \frac{1}{N_j^{(n)}} \sum_{k=1}^{N_j^{(n)}} T_j^{(k)}}{\sum_i \frac{N_i^{(n)}}{n} \frac{1}{N_i^{(n)}} \sum_{k=1}^{N_i^{(n)}} T_i^{(k)}}.$$

Taking the limit in the above equation and applying the strong law of large numbers, we obtain

$$f_j = \frac{\pi_j E(T_j^{(1)})}{\sum_i \pi_i E(T_i^{(1)})} = \frac{\pi_j \alpha_j}{\sum_{i\in E} \pi_i \alpha_i},$$

verifying the assertion (54). □

6.2. Limit distributions for mirror semi-Markov processes. Let $Z = (X, Y)$ be a time homogeneous mirror semi-Markov process with a discrete state space E. We assume that $\zeta := \sum_{k\geq 0} Y_k = \infty$. It can be uniquely determined by the following regime:

$$X_0 \xrightarrow{Y_0} X_1 \xrightarrow{Y_1} \cdots X_n \xrightarrow{Y_n} \cdots.$$

We now let $\hat{E}' = \{(i, j); P_{ij} > 0\}$, $\hat{X} = \{\hat{X}_n = (X_n, X_{n+1}); n \geq 0\}$ and $\hat{Y} = \{Y_{n+1}; n \geq 0\}$. Then $\hat{Z} = (\hat{X}, \hat{Y})$ is a simple WMSP with state space $\hat{E}'$, and is a special case of a semi-Markov process. It can be described by the following regime:

$$(X_0, X_1) \xrightarrow{Y_1} (X_1, X_2) \xrightarrow{Y_2} \cdots (X_n, X_{n+1}) \xrightarrow{Y_{n+1}} \cdots.$$

Use the notation $\hat{N}_t = \max\{n \geq 0; \sum_{k=1}^n Y_k \leq t\}$ (with the convention $\max\emptyset = 0$). Then the process $\hat{Z}_t$ defined by (5) and (6) with (X, Y) replaced by $(\hat{X}, \hat{Y})$ can be expressed as $\hat{Z}_t = \hat{X}_{\hat{N}_t}$. In what follows the corresponding notations related to $\hat{Z}$ will be marked by the superscript ˆ.

THEOREM 6.3. *Suppose that the embedded Markov chain X admits a unique stationary distribution $\Pi = (\pi_i)_{i\in E}$ with $\pi_i > 0$ for all $i \in E$, and the initial distribution of X_0 is Π. Let $j \in E$. If $\hat{V}^{(1)}_{(i,j)(i,j)}$ is not a lattice random variable and $\hat{\mu}_{(i,j)(i,j)} = E\hat{V}^{(1)}_{(i,j)(i,j)} < \infty$ for all $(i, j) \in \hat{E}'$. Then $f_j := \lim_{t\to\infty} P(Z(t) = j)$ exists and*

$$f_j = \frac{\alpha_j}{\mu_{jj}} = \frac{\alpha_j \pi_j}{\sum_{i\in E} \alpha_i \pi_i} = \lim_{t\to\infty} \frac{\int_0^t 1_{\{Z_s = j\}} \mathrm{d}s}{t}, \quad a.s. \tag{55}$$

PROOF. By the assumption of the theorem we can verify that $\hat{X}$ is irreducible and recurrent. Applying Theorem 6.2, we have

$$\hat{f}_{(i,j)} := \lim_{t\to\infty} P\big(\hat{Z}(t) = (i,j)\big) = \frac{\hat{\alpha}_{(i,j)}}{\hat{\mu}_{(i,j)(i,j)}} = \lim_{t\to\infty} \frac{\int_0^t 1_{\{\hat{Z}_s=(i,j)\}}\mathrm{d}s}{t}, \quad a.s.$$

Then

$$\begin{aligned}
\lim_{t\to\infty} P\big(Z(Y_0+t) = j\big) &= \lim_{t\to\infty} P\Big(\bigcup_{i\in E} \hat{Z}(t) = (i,j)\Big) \\
&= \lim_{t\to\infty} \sum_{i\in E} P\big(\hat{Z}(t) = (i,j)\big) \\
&= \sum_{i\in E} \lim_{t\to\infty} \frac{\int_0^t 1_{\{\hat{Z}_s=(i,j)\}}\mathrm{d}s}{t}, \quad a.s. \\
&= \lim_{t\to\infty} \frac{\int_0^t 1_{\{\tilde{Z}_s^{(1)}=j\}}\mathrm{d}s}{t}, \quad a.s.
\end{aligned}$$

Employing the notations and results of Theorem 5.11 in Section 5, we have

$$\begin{aligned}
\hat{\mu}_{(i,j)(i,j)} &= \frac{\sum_{(m,n)\in\hat{E}'} \hat{\alpha}_{(m,n)}\hat{\pi}_{(m,n)}}{\hat{\pi}_{(i,j)}} \\
&= \frac{\sum_{(m,n)\in\hat{E}'} {}_m\alpha_n \cdot \pi_m \cdot P_{mn}}{\pi_i P_{ij}} \\
&= \frac{\sum_{(m,n)\in E\times E} {}_m\alpha_n \cdot \pi_n \cdot c_{mn}}{\pi_j c_{ij}} \\
&= \frac{\sum_{n\in E} \alpha_n\pi_n}{\pi_j c_{ij}} \\
&= \frac{\mu_{jj}}{c_{ij}}.
\end{aligned}$$

Therefore

$$\begin{aligned}
\lim_{t\to\infty} P\big(Z(Y_0+t) = j\big) &= \lim_{t\to\infty} P\Big(\bigcup_{i\in E} \hat{Z}(t) = (i,j)\Big) \\
&= \lim_{t\to\infty} \sum_{i\in E} P\big(\hat{Z}(t) = (i,j)\big) \\
&= \sum_{i\in E} \frac{\hat{\alpha}_{(i,j)}}{\hat{\mu}_{(i,j)(i,j)}} \\
&= \sum_{i\in E} \frac{c_{ij}\cdot\hat{\alpha}_{(i,j)}}{c_{ij}\cdot\hat{\mu}_{(i,j)(i,j)}} \\
&= \sum_{i\in E} \frac{c_{ij}\cdot {}_i\alpha_j}{\mu_{jj}}
\end{aligned}$$

$$= \frac{\alpha_j}{\mu_{jj}} = \frac{\alpha_j \pi_j}{\sum_{i\in E} \alpha_i \pi_i}.$$

Hence we get

$$\lim_{t\to\infty} P\big(Z(Y_0+t)=j\big) = \frac{\alpha_j}{\mu_{jj}} = \frac{\alpha_j\pi_j}{\sum_{i\in E}\alpha_i\pi_i} = \lim_{t\to\infty} \frac{\int_0^t 1_{\{Z_{Y_0+s}=j\}}\mathrm{d}s}{t}, \quad a.s.$$

Due to the fact that $Y_0 < \infty$ a.s., we conclude that $f_j := \lim_{t\to\infty} P(Z(t)=j)$ exists and

$$f_j = \frac{\alpha_j}{\mu_{jj}} = \frac{\alpha_j\pi_j}{\sum_{i\in E}\alpha_i\pi_i} = \lim_{t\to\infty} \frac{\int_0^t 1_{\{Z_s=j\}}\mathrm{d}s}{t}, \quad a.s. \qquad \square$$

7. A remark on simple WMSP. A key step in proving Theorem 6.3 is to reduce the mirror semi-Markov process into a simple WMSP. This idea is powerful and can be applied to more general cases, which reveals that the class of simple WMSPs is of importance in theoretical study.

Suppose that $Z = (X, Y)$ is a time homogeneous (after n_0) WMSP with a discrete state space E. Let S_n, S_n^+, S_n^-, d, d^+ and d^- be the corresponding notations specified in Section 4. If $S_n = \{n-d^-, \dots, n-1, n, n+1, \dots, n+d^+\}$ for each $n \ge n_0$, then we can use the following method to reduce Z into a simple WMSP.

We define $\hat{E} = \{(i_1,\dots,i_d); i_1,\dots,i_d \in E\}$. For $(i_1,\dots,i_d) \in \hat{E}$ and $(j_1,\dots,j_d) \in \hat{E}$, we define

$$\hat{P}\big((i_1,\dots,i_d),(j_1,\dots,j_d)\big) = \begin{cases} P_{i_d j_d} & \text{if } i_k = j_{k-1}, \quad \text{for all } k = 2,\dots,d, \\ 0 & \text{otherwise}, \end{cases}$$

where P_{ij} is the original transition probability of X. We set

$$\hat{X}_n = (X_{n+n_0-d^-}, \dots, X_{n+n_0-1}, X_{n+n_0}, X_{n+n_0+1}, \dots, X_{n+n_0+d^+}), \quad \text{for all } n \ge 0.$$

Then the process $\hat{X} = \{\hat{X}_n, n \ge 0\}$ is a Markov chain on the state space $\hat{E}$ with the transition probability:

$$\hat{P} = \big\{\hat{P}\big((i_1,\dots,i_d),(j_1,\dots,j_d)\big)\big\}.$$

We write $\hat{Y} = \{Y_{n+n_0}, n \ge 0\}$. Then the new process $\hat{Z} = (\hat{X}, \hat{Y})$ is a simple WMSP with the state space $\hat{E}$. Moreover, $\hat{Z}$ inherits time homogeneity from Z. By this method, in some cases we can borrow the results of simple WMSP for the study of a time homogeneous WMSP, as we have done in the proof of Theorem 6.3. But we should aware that during the course of inducing a general WMSP Z into a simple WMSP $\hat{Z}$, we may also lose some good properties of the original WMSP. For example, suppose that the embedded Markov chain X of the original Z is irreducible, and that there exists a pair of states $\{i, j\} \subset E$ such that $P_{ij} = 0$. Then the new Markov chain $\hat{X}$ will never reach the states $(j_1,\dots,j_d)$ of $\hat{E}$ which contain the ordered pair $\{i, j\}$, and hence the new embedded Markov chain $\hat{X}$ will become reducible.

8. Applications to Web page ranking. In this section we explore how the framework of WMSP can be applied to the study of web page ranking. Recently the authors in [9, 10] found that the page importance is mainly affected by two factors: the page reachability, the average possibility that the surfer arrives at the page; and the page utility, the average value that the page gives to the surfer in a single visit. The page reachability can be quantized as the value of the stationary distribution of the embedded Markov chain, and the page utility can be quantized as the mean staying time on a page. Thus the page importance can be represented as the product of these two trackable factors. We show that the framework of WMSP is very suitable to analyze the page importance in this way. The framework can cover many existing algorithms including PageRank, TrustRank, and BrowseRank as its special cases. We show also that the framework can help us design new algorithms to handle more complex problems. For example, we can employ mirror semi-Markov processes to design MobileRank, handling mobile Web which differs a lot from the usual Web structurally.

8.1. PageRank and discrete time Markov processes. PageRank is one of the most famous link analysis algorithms for the page importance ranking. It was proposed by Brin and Page in 1998 [4, 27], and has been successfully used by the Google search engine. In this algorithm, a PageRank matrix is constructed by the link graph, and the surfing on the Web is modeled as a discrete time Markov chain with PageRank matrix as its transition probability matrix. It employs a power method [11] to calculate the stationary distribution of the Markov chain. Accordingly the stationary distribution is interpreted as the PageRank values of web pages (see Langville and Meyer [18]). Here we briefly describe the discrete time Markov process employed in PageRank.

We regard the hyperlink structure of web pages on a network as a directed graph $\tilde{G} = (\tilde{V}, E)$. A vertex $i \in \tilde{V}$ of the graph represents a web page, and a directed edge $\overrightarrow{ij} \in E$ represents a hyperlink from the page i to the page j. Suppose that $|\tilde{V}| = N$. Let $\tilde{B} = (\tilde{B}_{ij})_{N\times N}$ be the adjacent matrix of $\tilde{G}$ and $\tilde{b}_i$ be the sum of the i-th row of $\tilde{B}$, i.e. $\tilde{b}_i = \sum_{j=1}^{N} \tilde{B}_{ij}$. If $\tilde{b}_i = 0$, then we let all entries of the i-th row of B equal to 1. By this way we get a modified matrix B. Denote by b_i the sum of the i-th row of B and by D the diagonal matrix with diagonal entry b_i. Now, we construct a stochastic matrix $\bar{P} = D^{-1}B$.

When a surfer browses on the Web, he may choose the next page by randomly clicking one of the hyperlinks in the current page with a large probability α (in practice it is often set $\alpha = 0.85$), which means that with probability α, the surfer may randomly walk on $\tilde{G}$ with transition probability $\bar{P}$; while with probability $(1-\alpha)$, the surfer may also open a new page from the Web, and the new page might be selected randomly according to his personal preference φ, which means that he walks randomly on $\tilde{G}$ with transition probability $e^T\varphi$, where e is a row vector of all ones, and φ is an N-dimensional probability vector (in practice it is often set $\varphi = e/N$ for simplicity), which is called the personalized vector. Combining the above two random walks, an irreducible transition matrix, the so called PageRank matrix, is formulated as follows,

$$P = \alpha\bar{P} + (1-\alpha)e^T\varphi\,. \tag{56}$$

The browsing behavior of a random surfer is then modeled as a discrete time Markov process X with the transition matrix P specified by (56). As has being illustrated in Example 1 of Section 3, it is a special WMSP. In the framework of WMSP, we may write it as $Z = (X, Y)$ with $Y_n \equiv 1$ for all n.

In practice, researchers assume usually that the Markov chain X is aperiodic due to the huge number of web pages. Then by the ergodicity theorem, X admits a unique stationary distribution $\Pi = (\pi_1, \pi_2, \ldots, \pi_N)$, which satisfies

$$\Pi = \Pi P\,, \quad \text{and} \quad \sum\nolimits_{k=1}^{N} \pi_k = 1\,. \tag{57}$$

In PageRank, the page reachability is computed as the above stationary distribution Π, and the page utility is equal to one for all pages under the assumption that all pages are equally useful for the random surfer, which is indicated by $P(Y_n = 1 \mid \mathscr{F}^X) = 1,$ for all $n \geq 0$.

As TrustRank is a modified PageRank by starting the iterative process from a reliable seed set, it can also be covered by the framework of WMSP. Similarly, other models of PageRank-alike algorithms can be regarded as special cases of WMSPs.

8.2. BrowseRank family and simple WMSPs. Although PageRank has many advantages, recently people have realized that it has also certain limitations as a model for representing page importance. For example, the link graph, which PageRank relies on, is not a very reliable data source, because hyperlinks on the Web can be easily added or deleted by the web content creators. It is clear that those purposely created hyperlinks (e.g. created by link farms) are not suitable for calculating the page importance. To tackle the limitations of PageRank, recently a family of new algorithms called the BrowseRank family has been proposed [19, 20, 21].

The BrowseRank algorithm was first introduced in [19], where the browsing behavior of a random surfer is modeled as a time homogeneous continuous-time Markov process (i.e. Q-process) $\{Z_t, t \geq 0\}$, which is called the user browsing process. By the algorithm one should collect the user behavior data in web surfing and build a user browsing graph, which is different from the traditional link graph, refer to [22] for more details. The process takes its values in the state space V consisting of all the web pages in the user browsing graph. Thus the user browsing process Z contains both user transition information and user staying time information. The evaluation of Z_t at time t represents the web page that the random surfer is browsing at the time point t, where t may take value in the set $\boldsymbol{R}^+$ of all the nonnegative real numbers. Afterwards, a number of variations of the algorithm, which we refer to as the BrowseRank family, were introduced and discussed in [20, 21]. The browsing processes modeled in the BrowseRank family are not necessarily Q-processes, but they are all time homogeneous simple WMSPs as described by Example 4.5 (A) in Section 4.

For a time homogeneous simple WMSP $Z = (X, Y)$, the conditional distribution of the random variable Y_n depends only on X_n, which means that the length of staying time on a page depends only on the current page itself. Moreover, by the time homogeneity the conditional distribution is independent of n. We introduce a random variable T_i representing the staying time on the state i. Then the mean staying time on the page i can be calculated by the following formula,

$$\alpha_i := E(T_i) = \int_0^\infty t F_i(\mathrm{d}t) .$$

For a browsing process $Z = (X, Y)$, the embedded Markov chain X is always constructed to be irreducible and recurrent. Then we can calculate the page importance of web pages by Theorem 6.2. In particular, if the browsing process Z is a Q-process, then the distribution function of T_i is given by $F_i(t) = 1 - e^{-\lambda_i t}$, for all $t \geq 0$. Here λ_i is a parameter which depends on the page i. In this case the mean staying time on the page i can be explicitly calculated as follows,

$$\alpha_i = E(T_i) = \int_0^\infty t F_i(\mathrm{d}t) = \int_0^\infty t\lambda_i e^{-\lambda_i t}\, \mathrm{d}t = \frac{1}{\lambda_i} .$$

Let Π be the the unique stationary distribution of X determined by (57). Then the page importance, which was called the BrowseRank score in [21], can be calculated by (54) as follows,

$$f_i = \theta \frac{\pi_i}{\lambda_i} , \quad \text{for all } \ i \in E ,$$

where $\theta = \big(\sum_{j \in E} \pi_j / \lambda_j\big)^{-1}$.

In fact in [21] the authors discussed 8 variations of BrowseRank algorithms family. Among them 4 variations ($BR(M, \cdot)$ family) are based on simple WMSPs, and 4 variations ($BR(A, \cdot)$ family) are based on Q-processes.

8.3. ExtBrowseRank and semi-Markov processes. In the algorithms mentioned above, the staying time on a page is assumed to depend only on the current page and to be independent of other pages. As a first approximation this assumption is acceptable. But if we want to have a more accurate model, we will find that this assumption is too rough. Actually, when a user is browsing on web pages to search some desired information, sometimes the staying time will not depend only on the current page, but depends also on the next page. For example, if he finds a hyperlink directing to a very interesting page, or he is attracted by a beautiful advertisement displayed on the page, he may immediately jump to the next more attractive page, without completing his current reading. In this situation we may employ a semi-Markov process to model the browsing behavior of the user.

Let $Z = (X, Y)$ be a semi-Markov process which is employed to model a browsing behavior. The (conditional) distribution of the random variable Y_n depends on two successive states X_n and X_{n+1}, which means that the length of the n-th staying time depends not only on the current page, but also on the next page which the user will visit. Suppose that Z is time homogeneous. Then the distribution will be independent of n. We may introduce a random variable T_{ij} representing the staying time on the state i by knowing the next state j. We introduce also a random variable T_i representing the staying time on the state i without knowing the next state. By (35) and (36) we have:

$$F_{ij}(t) = P(T_{ij} \le t) = P(Y_n \le t \mid X_n = i, X_{n+1} = j) = \frac{Q_{ij}(t)}{P_{ij}}.$$

By the formula of the total probability we have,

$$F_i(t) := P(T_i \le t) = \sum_{j \in E} F_{ij}(t) P_{ij}. \tag{58}$$

Similarly as the BrowseRank algorithm, we define the page reachability as the stationary distribution of the embedded Markov chain X if it exists, and define the page utility as the mean staying time on the considered page. Therefore, the page importance can be calculated by equations (53) and (54). The stationary distribution Π involved in (54) can be determined by (57). To compute the mean staying times, we may design the following extension of the Browserank algorithm, which we call the ExtBrowseRank algorithm.

We denote by α_{ij} the expectation of T_{ij}.

Step 1: Suppose that the transition probability matrix P of X, and the time information α_{ij}, for all $i, j \in E$ have been estimated from real users browsing log data.

Step 2: Estimate the distribution function of staying time on page i, for all $i \in E$, by (58) mentioned above.

Step 3: Compute the mean staying time α_i, for all $i \in E$, by (59) below.

$$\begin{aligned} \alpha_i &= \int_0^\infty t F_i(\mathrm{d}t) \\ &= \sum_{j \in E} \int_0^\infty t P_{ij} F_{ij}(\mathrm{d}t) \\ &= \sum_{j \in E} P_{ij} \alpha_{ij}. \end{aligned} \tag{59}$$

Step 4: Compute the page importance scores according to (60) below, which is deduced also from (54).

$$f_i = \frac{\pi_i \sum_{j \in E} P_{ij} \alpha_{ij}}{\sum_{i \in E} \left(\pi_i \sum_{j \in E} P_{ij} \alpha_{ij}\right)}. \tag{60}$$

8.4. MobileRank and mirror semi-Markov processes. In recent years a new type of Web structure, namely mobile Web structure, developed rapidly. Mobile Web differs largely from general Web in several aspects. For example, the topology of the mobile Web graph differs significantly from that of the general Web graph [17]. Because the owners of websites on mobile Web tend to create hyperlinks only to their own pages or pages of their business partners. Consequently there are more disconnected components in the mobile web graph, and links do not always mean recommendation, but often mean business connection. In this situation, it is clear that the scores computed by existing algorithms like PageRank may not reflect the true importance of the pages. By employing mirror semi-Markov processes, the authors in [9, 10] proposed a new algorithm called MobileRank for computing the page importance on mobile Web. Actually they found a new way of calculating mean staying time. Here we just briefly introduce part of the related work. For more details the reader is referred to [9, 10].

Let $Z = (X, Y)$ be a time homogeneous mirror semi-Markov process which is employed to model a browsing behavior on mobile Web. Then the staying time Y_n depends not only on the current page X_n, but also on the inlink page X_{n-1}. This means that the mirror semi-Markov process has the ability to make use of the inlink information for promoting or demoting staying time on the current page. For example, if an inlink is from a partner website, then the user may demote the staying time of visits from the website. Introduce a random variable ${}_kT_i$ representing the staying time on page i by knowing the previous page k (cf. the beginning of Subsection 5.3). By (37) and (38) we have

$$ {}_kF_i(t) := P({}_kT_i \le t) = P(Y_n \le t | X_{n-1} = k, X_n = i) = \frac{{}_kQ_{ij}(t)}{P_{ij}} . $$

Suppose that the assumption of Lemma 5.9 is fulfilled. By (51) and (50) we obtain

$$ F_i(t) = \sum_{k \in E} {}_kF_i(t) c_{ki} = \frac{\sum_{k \in E} {}_kF_i(t) P_{ki} \pi_k}{\pi_i} , $$

where $\{c_{ki}\}$ is the contribution probability introduced in Subsection 5.3 and $\{\pi_i\}$ is the stationary distribution of X. Then we can calculate the expectation of $F_i(t)$, which is called *the mean staying time on the state* i in [9], by the formula below,

$$ \begin{aligned} \alpha_i &= \int_0^\infty t F_i(\mathrm{d}t) \\ &= \sum_{k \in E} \frac{P_{ki}\pi_k}{\pi_i} \int_0^\infty t_k F_i(\mathrm{d}t) \\ &= \sum_{k \in E} \frac{{}_k\alpha_i P_{ki}\pi_k}{\pi_i} . \end{aligned} $$

Further we can calculate the limit distribution $f_i := \lim_{t \to \infty} P(Z(t) = i)$ of Z by the formula (55) as follows,

$$ f_i = \frac{\sum_{k \in E} {}_k\alpha_i P_{ki}\pi_k}{\sum_{i \in E} \left(\sum_{k \in E} {}_k\alpha_i P_{ki}\pi_k \right)} , \quad \text{for all } i \in E . $$

By Theorem 6.3 we know that f_i is the proportion of the average sojourn time on the page i. Therefore it was used in [9] as the page importance score.

Acknowledgments. We are grateful to Bin Gao, Tie-Yan Liu and Hang Li from MSRA for stimulating discussions and valuable suggestions during our pleasant collaborations. The work is partly supported by National Natural Science Foundation of China (Grant No. 11001010), China Postdoctoral Science Foundation funded project (Grant No. 20100470198), the Fundamental Research Funds for the Central Universities (Grant No. 2009JBM106), 973 project, and the key lab project of CAS. The authors thank the referee and the editorial office for their very helpful comments and suggestions.

REFERENCES

[1] Y. BAO, G. FENG, T. LIU, Z. M. MA AND Y. WANG, Ranking websites, a probabilistic view, Internet Math. 3 (2007), 295–320.
[2] Y. BAO, Z. M. MA AND Y. SHANG, Some recent results on ranking webpages and websites, in Industrial and Applied Mathematics, series in comtemporary applied mathematics CAM 10, Higher Education Press, Beijing, 2009.
[3] V. BARBU AND N. LIMNIOS, Semi-Markov chains and hidden semi-Markov models towards applications, Lecture Notes in Statist. 191, Springer, New York, 2008.
[4] S. BRIN AND L. PAGE, The anatomy of a large-scale hypertextual Web search engine, Computer Networks and ISDN Systems 30 (1998), 107–117.
[5] D. CAI, X. HE, J. WEN AND W. MA, Block-level link analysis, In proceedings of SIGIR '2004, 440–447.
[6] E. ÇINLAR, Markov renewal theory: A Survey, Management Science 21 (1975), 727–752.
[7] Y. DENG AND Z. LIANG, Random point process and its application (in Chinese), Science Press, 1992.
[8] G. FENG, T. LIU, Y. WANG, Y. BAO, Z. M. MA, X. D. ZHANG AND W. MA, AggregateRank: Bringing Order to Web Sites, In proceedings of SIGIR '2006, 75–82.
[9] B. GAO, T. LIU, Y. LIU, T. WANG, H. LI AND Z. M. MA, Page importance computation based on Markov processes, Information Retrieval 14 (2011), 488–514.
[10] B. GAO, T. LIU, Z. M. MA, T. WANG AND H. LI, A general Markov framework for page importance computation, In proceedings of CIKM '2009, 1835–1838.
[11] G. H. GOLUB AND C. F. V. LOAN, Matrix computations (3rd ed.), Johns Hopkins University Press, 1996.
[12] Z. GYONGYI, H. GARCIA-MOLINA AND J. PEDERSEN, Combating web spam with trustrank, In proceedings of VLDB '2004, 576–587.
[13] Z. HOU, W. LIU, Z. LIU, G. LIU, J. ZHOU, J. LI AND C. YUAN, Markov skeleton processes, Hunan Science and Technology Press, 2000.
[14] Z. HOU AND G. LIU, Markov skeleton processes and their applications, Science Press and International Press, 2005.
[15] J. JACOD, Calcul stochastigue et probléḿs de martingales, Lecture Notes in Mathematics 714, Springer, 1979.
[16] J. JACOD, Multivariate point process: Predictable projection, Radon-Nikodym Derivative, Representation of Martingales, Probab. Theory Relat. Fields 31 (1975), 235–253.
[17] A. JINDAL, C. CRUTCHFIELD, S. GOEL, R. KOLLURI AND R. JAIN, The mobile web is structurally different, In proceedings of the 11th IEEE Global Internet Symposium' 2008.
[18] A. LANGVILLE AND C. MEYER, Deeper inside pagerank, Internet Math. 1 (2004), 335–400.
[19] Y. LIU, B. GAO, T. LIU, Y. ZHANG, Z. M. MA, S. HE AND H. LI, BrowseRank: Letting users vote for page importance, In proceedings of SIGIR '2008, 451–458.
[20] Y. LIU, Research on stochastic models and algorithms for web page ranking (in Chinese), Doctor degree thesis, 2009.
[21] Y. LIU, T. LIU, B. GAO, Z. M. MA AND H. LI, A framework to compute page importance based on user behaviors, Information Retrieval 13 (2010), 22–45.
[22] Y. LIU AND Z. M. MA, Comparison of two algorithms for computing page importance, AAIM2010, LNCS 6124 (2010), 1–11.
[23] Y. LIU, Z. M. MA AND C. ZHOU, Further study on web Markov skeleton processes (in preparation).
[24] Z. M. MA AND L. TANG, Martingale method to minimum Q-process (in Chinese), Hunan Ann. Math. 3 (1983), 36–44.
[25] U. N. BHAT, Elements of applied stochastic processes (2nd edition), John Wiley & Sons, Inc., 1984.
[26] Z. NIE, Y. ZHANG, J. WEN AND W. MA, Object-level ranking: Bringing order to web objects, In proceedings of WWW '2005, 567–574.
[27] L. PAGE, S. BRIN, R. MOTWANI AND T. WINOGRAD, The pagerank citation ranking: Bringing order to the

web, Technical report, Stanford Digital Library Technologies Project, 1998.

[28] S. HE, J. WANG AND J. YAN, Semimartingale theory and stochastic calculus (Science Press), Beijing CRC Press, Boca Raton, FL, 1992.

SCHOOL OF SCIENCE
BEIJING JIAOTONG UNIVERSITY
NO. 3, SHANGYUAN RESIDENCE
BEIJING, 100044
CHINA

E-mail address: ytliu@bjtu.edu.cn

ACADEMY OF MATHEMATICS AND
SYSTEMS SCIENCE
CHINESE ACADEMY OF SCIENCES
NO. 55, ZHONGGUANCUN EAST ROAD
BEIJING, 100190
CHINA

E-mail addresses: mazm@amt.ac.cn
: zhouchuan@amss.ac.cn