2025 IEEE 5th International Conference on Smart Information Systems and Technologies (SIST)
14–16 May, 2025, Astana, Kazakhstan

2025 IEEE 5th International Conference on Smart Information Systems and Technologies (SIST) | 979-8-3315-1514-0/25/$31.00 ©2025 IEEE | DOI: 10.1109/SIST61657.2025.11139228

Data-Driven Markovian Project Portfolio Tracking

Thomas A. Weber
École Polytechnique Fédérale de Lausanne
Lausanne, Switzerland
thomas.weber@epfl.ch

***Abstract*—We propose a finite-state Markov chain framework for tracking and forecasting the status of project portfolios. This approach enables forecasts of portfolio composition over time and the computation of long-run distributions of project outcomes. It supports strategic planning by identifying project success rates, average durations, and the balance of resource allocation between active and idle projects. From a managerial perspective, the model facilitates early detection of portfolio-level risks and provides a data-driven basis for adjusting resource deployment or re-prioritizing projects. We show that forecasts remain robust under moderate errors in model identification, enhancing the method's practical applicability in environments with noisy or incomplete data. This work lays the foundation for a scalable, organization-wide mechanism to improve visibility into project dynamics and support evidence-based decision-making.**

***Keywords*—Project portfolio tracking, Markov chains, project management, phase-gate approach.**

I. Introduction

Managing long-term projects in industries such as pharmaceuticals requires balancing strategic objectives with operational realities. Large-scale IT and business transformation initiatives, in particular, involve multiple interdependent projects with varying degrees of complexity, uncertainty, and risk. Maintaining visibility into the evolving status of a project portfolio is critical for decision-making, resource allocation, and risk mitigation. Traditional qualitative assessments, such as periodic reporting and expert judgment, while available, are usually not aggregated for predictive purposes or for evaluating portfolio performance across relevant dimensions, such as success rate, project duration, or resource balance. Such aggregate measures can also serve as benchmarks for individual project performance, for example, by indicating their positions relative to group averages.

Here, we propose a finite-state Markov chain approach to project portfolio tracking. By modeling project states using a structured framework, this method enables quantitative forecasting of future portfolio status over finite time horizons and provides insights into the long-run distribution of project outcomes. The structured transition dynamics offer an objective foundation for tracking progress and adapting strategies in response to evolving risks. Beyond forecasting, an important concern in project management is the robustness of model-based assessments. Misidentifying initial project states or transition probabilities can introduce errors into portfolio projections, potentially leading to suboptimal decisions. To evaluate the impact of such errors, we analyze how identification inaccuracies propagate through the model.

The systematic management of projects dates back over a century to the structured bar-chart and harmonogram approaches of Gantt [1] and Adamiecki [2], respectively. Managing project portfolios as groups of projects originates in Markowitz's work on portfolio investment [3], [4]. Large-scale industrial and public projects in the 1950s and 1960s, including NASA initiatives, necessitated phased project planning, leading to the phase-gate approach popularized by Cooper [5], [6], followed by Agile frameworks that allow for dynamic adjustments instead of frontloading [7]. The persistent monitoring of projects remains an ongoing concern in most large organizations [8].[1]

The motivation for this research stems from various industry projects in the pharmaceutical sector, where companies have struggled to forecast the state of their project portfolios, particularly those involving IT projects [10]. The core idea is to implement a uniform classification system across the organization so as to leverage past experience from observed state transitions, thus enabling data-driven forecasts based on Markov chains; see, e.g., Feller [11, Ch. XV]. Thus far, Markov chains have been applied primarily to individual projects—for example, in the context of Markov PERT networks to estimate the duration of a single project [12], [13].[2] Our contribution lies in adapting Markov chain modeling to the dynamic tracking of project portfolios, providing a structured, data-driven methodology for anticipating portfolio evolution, evaluating its performance, and mitigating risks.

II. Model

Consider a firm with a portfolio of projects. Each project is either *active* or *passive*. As long as a project is active, it is assessed at regularly spaced (integer-valued) times t for its current *state*, which is an element of the (nonempty) finite *state space* $\mathcal{S}$. While our framework applies to any finite state space, in order to fix ideas, we assume that the firm uses a "traffic-light approach" to assess active projects, so

$$\mathcal{S} = \{\text{Green } (G), \text{Yellow } (Y), \text{Red } (R)\}\,.$$

The key performance indicators (KPIs) used for the periodic evaluation of active projects and their aggregation into scores

[1]For a survey of project portfolio management since 1950, see [9].
[2]PERT: Program Evaluation and Review Technique [14].

with classification thresholds may vary from one organization to another. The methods could also vary with the nature of the projects, such as whether they deal primarily with cost effectiveness, human resources, information technology, innovation, marketing, or another particular field (or combination thereof). Important for the use of our proposed method is not necessarily that projects all have to come from the same domain or be very similar to each other (e.g., in size or duration), but rather that the assessment methods are deemed comparable.

Example 1. In the traffic-light approach, the project status may be assessed as follows:

- Green: On track, within budget and timeline.
- Yellow: At risk, some issues emerging.
- Red: Critical, major delays or budget overruns.

A passive (or *inactive*) project can be in one of three service states: New (N), Completed (C), or Discontinued (D). Before becoming an active project it passes through a planning and preparation phase which coincides with the service state N. Upon exiting this preparation phase, the project becomes active until it is *either* completed (entering the service state C) *or* discontinued (i.e., passing into the service state D).

Remark 1. The total number of states any (active or passive) project can be in is $n = |\mathcal{S}| + 3 \geq 4$. Under the traffic-light approach, it is $n = 6$.

A. *Project Portfolio Markov Chain*

The current *portfolio distribution* is captured by a vector $\boldsymbol{p} = (p_1, \ldots, p_n)$ that contains the proportions of projects in each of n possible states $s_1, \ldots, s_n$ (either in $\mathcal{S}$ for active projects, or in one of the three service states). Thus, it is an element of the $(n-1)$-dimensional domain of discrete probability distributions,

$$\Delta_n = \{(\omega_1, \ldots, \omega_n) \in \mathbb{R}^n_+ : \omega_1 + \cdots + \omega_n = 1\}.$$

From one time period to the next, the state of any given project transitions from its current state s_i to another state s_j with probability $p_{ij} \in [0, 1]$, where $i, j \in \mathcal{N} = \{1, \ldots, n\}$.

Example 2. In the traffic-light approach, we set $(N, G, Y, R, C, D) = (1, 2, 3, 4, 5, 6)$ to denote each state interchangeably by its index (in $\mathcal{N}$) or by its letter specification (in $\{N, G, Y, R, C, D\} = \mathcal{N}$). For example, $s_N = s_1$, or $s_R = s_4$.

The matrix of transition probabilities $\mathbf{P} = [p_{ij}] \in \mathbb{R}^{n \times n}$ is stochastic, that is, it has the 'Markov property' in the sense that each of its rows sums to one,

$$\sum_{j=1}^{n} p_{ij} = 1, \quad i \in \mathcal{N},$$

as any given project must transition from one of the n feasible states to another one of those states with probability 1. Fig. 1 shows the Project Portfolio Markov Chain with its primitives (set of states and transition probabilities) under the traffic-light approach.

Example 3. A possible transition matrix for this six-state system in Ex. 2 is given by:

$$\mathbf{P} = \begin{bmatrix} p_{NN} & p_{NG} & 0 & 0 & 0 & p_{ND} \\ 0 & p_{GG} & p_{GY} & p_{GR} & p_{GC} & p_{GD} \\ 0 & p_{YG} & p_{YY} & p_{YR} & p_{YC} & p_{YD} \\ 0 & p_{RG} & p_{RY} & p_{RR} & p_{RC} & p_{RD} \\ p_{CN} & 0 & 0 & 0 & p_{CC} & 0 \\ p_{DN} & 0 & 0 & 0 & 0 & p_{DD} \end{bmatrix}.$$

The first and the last two rows of $\mathbf{P}$ correspond to the service states, while the other rows capture transitions from active project states in $\mathcal{S}$. If $p_{CN} = p_{DN} = 0$, then project completion (C) and discontinuation (D) become absorbing states,[3] for any project to remain there permanently.

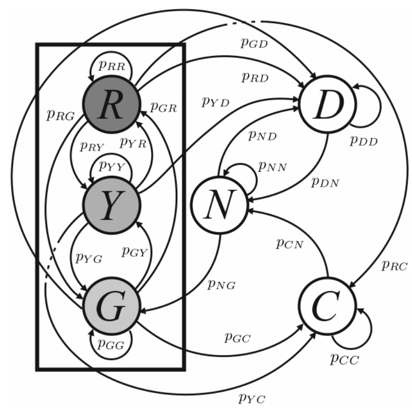

Fig. 1. Project portfolio Markov chain with traffic-light states and service states, with index set $\mathcal{N} = \{N, G, Y, R, C, D\}$, as in Exs. 2 and 3.

B. *Model Identification*

To estimate the transition probabilities involving states in $\mathcal{S} \cup \{C, D\}$, the firm needs to analyze historical project data, based on review reports, answering questions such as:

- What is the probability that a Green project turns Yellow in the next review period?
- How likely is a Red project to recover to Yellow?
- What percentage of Yellow projects move to Green versus escalating to Red?

To determine the transition frequencies, it is necessary to:

- Count the number of transitions from one state to another;
- Normalize by the total occurrences of each initial state.

The transitions to and from state N involve the redeployment of project resources from completed or discontinued projects (via p_{CN} or p_{DN}, respectively), followed by continued planning (via p_{NN}), and subsequent release (via p_{NG}) of new projects into the active project pool. These transition probabilities can be estimated using longitudinal observations as explained in the following example.

Example 4. To identify p_{ij} when $i = N$ or $j = N$ (or both) in the transition matrix $\mathbf{P}$ of Ex. 3, one first observes the

[3] A state s_{ii} is called *absorbing* if $p_{ii} = 1$.

average residence time T_N of projects at the planning stage. For $p = p_{NN}$, this implies:

$$T_N = (1-p)\sum_{t=0}^{\infty} tp^t = (1-p)p\sum_{t=1}^{\infty} tp^{t-1}$$
$$= (1-p)p\,\frac{d}{dp}\sum_{t=1}^{\infty} p^t = \frac{p}{1-p},$$

by virtue of the geometric-series formula in the last step. Hence, one can identify

$$p_{NN} = \frac{T_N}{1+T_N} \quad \text{and} \quad p_{NG} + p_{ND} = 1 - p_{NN}.$$

For instance, if project planning and preparation (in a regime with a monthly project review) take on average one quarter of a year, then one would set $T_N = 3$ and therefore find that $p_{NN} = 0.75$. To disentangle p_{NG} and p_{ND}, let M be the number of projects that have passed through the planning stage over a sufficiently long time interval (greater than T_N), and of these it has been observed that only M_G have been successfully launched, with the remaining $M - M_G$ not launched, then $p_{ND}/p_{NG} = (M - M_G)/M_G$. Thus, in combination with the earlier relation one can conclude that

$$p_{NG} = \frac{M_G}{(1+T_N)\,M} \quad \text{and} \quad p_{ND} = \frac{M - M_G}{(1+T_N)\,M}.$$

Similarly, if T_C and T_D denote the respective average times until resources from completed and discontinued projects are redeployed, then by the same derivation as before, one finds that

$$p_{CC} = \frac{T_C}{1+T_C} \quad \text{and} \quad p_{DD} = \frac{T_D}{1+T_D}.$$

and consequently, $p_{CN} = 1 - p_{CC}$ and $p_{DN} = 1 - p_{DD}$. An example of a transition probability matrix for this six-state system is given by:

$$\mathbf{P} = \begin{bmatrix} 0.75 & 0.15 & 0 & 0 & 0 & 0.10 \\ 0 & 0.79 & 0.10 & 0.02 & 0.08 & 0.01 \\ 0 & 0.20 & 0.55 & 0.20 & 0.03 & 0.02 \\ 0 & 0.25 & 0.20 & 0.38 & 0.05 & 0.12 \\ 0.10 & 0 & 0 & 0 & 0.90 & 0 \\ 0.05 & 0 & 0 & 0 & 0 & 0.95 \end{bmatrix}.$$

The individual entries p_{ij} in $\mathbf{P}$ mean, for example, that projects in the *New* state remain there with 75% probability, transition to *Green* with 15% probability, or get *Discontinued* (10% of the time). Projects in the *Green* state stay in this state with 79% probability, transition to *Yellow* with 10% probability, to *Red* with 2% probability, get *Completed* (8%) or possibly also *Discontinued* (1%), and so forth.

III. Portfolio Evolution

The Project Portfolio Markov Chain (PPMC) discussed in Sec. II allows us to describe the evolution of project portfolio distributions, both forwards and backwards in time. We also establish the attenuation of errors in the project distribution data over time, providing for added extrapolation robustness, as a "regular" system (cf. Remark-2 below) evolves towards a steady-state distribution in the long run.

A. *Law of Motion*

Given a portfolio state $\boldsymbol{q} = (q_1, \ldots, q_n) \in \Delta_n$ at time zero, the evolution of the portfolio state from $\boldsymbol{q}$ to $\boldsymbol{q}^{(t)} = (q_1^{(t)}, \ldots, q_n^{(t)})$ after t periods proceeds according to the law of motion,[4]

$$\boldsymbol{q}^{(t)} = \boldsymbol{q}\,\mathbf{P}^t, \quad t \in \{0, 1, \ldots\}. \tag{1}$$

Equivalently, the probability of landing in state k after t periods starting from the portfolio distribution $\boldsymbol{q}$ is

$$q_k^{(t)} = \sum_{j=1}^{n} q_j p_{jk}^{(t)}, \quad (k,t) \in \mathcal{N} \times \{0, 1, \ldots\},$$

where the probability $p_{jk}^{(t)}$ of transitioning from s_j to s_k in exactly t periods can be computed as follows:

$$p_{jk}^{(t)} = \sum_{\nu=1}^{n} p_{j\nu} p_{\nu k}^{(t-1)}, \quad p_{jk}^{(0)} = \delta_{jk}, \quad t \geq 1,$$

for all $j, k \in \mathcal{N}$, where $\delta_{jk} = \mathbf{1}_{\{j=k\}}$ is the standard Kronecker symbol. In particular, it is $p_{jk}^{(1)} = p_{jk}$, and

$$\mathbf{P}^t = [p_{jk}^{(t)}], \quad t \geq 0.$$

Example 5. Given an estimated initial distribution $\boldsymbol{q} = (20\%, 30\%, 0\%, 10\%, 20\%, 20\%)$ and monthly review, the law of motion in Eq. (1) produces the 6-month forecast distribution $\boldsymbol{q}^{(6)} = \boldsymbol{q}\mathbf{P}^6 \approx (14.9\%, 24.0\%, 7.0\%, 3.1\%, 22.6\%, 28.4\%)$, where $\mathbf{P}$ is specified in Ex. 4.

Remark 2. Given $i \in \mathcal{N}$, a state s_i is *absorbing* if $p_{ii} = 1$. The PPMC is *irreducible* if every state can be reached (with positive probability) from every other state. If the PPMC has absorbing states (e.g., the states C and D can become absorbing if $p_{CC} = p_{DD} = 1$), it cannot be irreducible.

Let $f_{jk}^{(t)}$ be the probability that starting from the state s_j a project transitions *for the first time* to state s_k after exactly $t \geq 1$ steps. Then

$$p_{jk}^{(t)} = \sum_{\tau=1}^{t} f_{jk}^{(\tau)} p_{kk}^{(t-\tau)}, \quad j, k \in \mathcal{N}, \quad t \geq 1.$$

If we recall that $p_{kk}^{(0)} = 1$, then $f_{jk}^{(1)} = p_{jk}^{(1)}$, and more generally,

$$f_{jk}^{(t)} = p_{jk}^{(t)} - \sum_{\tau=1}^{t-1} f_{jk}^{(\tau)} p_{kk}^{(t-\tau)}, \quad t \geq 1.$$

Thus, $f_{jk} = \sum_{t=1}^{\infty} f_{jk}^{(t)}$ is the probability that the project currently in state s_j will *ever* reach state s_k. As long as the PPMC is irreducible, it is $f_{jk} = 1$ for all $j, k \in \mathcal{N}$ with $i \neq j$, and $(f_{jk}^{(t)})_{t \geq 1}$ describes the *first-passage distribution* for the state s_k conditional on the starting state s_j.

[4]It is $\boldsymbol{q}^{(0)} = \boldsymbol{q}$.

B. Steady-State Distribution

The steady-state (or *invariant*) distribution $\boldsymbol{\pi} = (\pi_1, \ldots, \pi_n)$ in Δ_n of an irreducible PPMC describes the long-run probabilities of a project being in each of its n possible states. Stationarity (so $\boldsymbol{\pi}^{(t)} \equiv \boldsymbol{\pi}$) requires by Eq. (1) that[5]

$$\boldsymbol{\pi}\mathbf{P} = \boldsymbol{\pi}, \tag{2}$$

together with the normalization condition,

$$\sum_{i=1}^{n} \pi_i = 1. \tag{3}$$

This distribution provides insights into the expected steady-state proportions of projects in each phase, enabling better resource allocation and strategic planning.

Remark 3. A *regular* PPMC is such that it is irreducible and aperiodic,[6] which is equivalent to $\mathbf{P}^t$ having only positive entries for some finite t [15, Thm. 4.1.2]. For example, given $\mathbf{P}$ as in Ex. 4 this is satisfied for $t \geq 3$. Any regular PPMC has a steady-state distribution which assigns positive weight on all states, so $\pi_j > 0$ for all $j \in \mathcal{N}$. When the PPMC contains absorbing states (so it is no longer regular), then the steady-state distribution assigns zero probability on all nonabsorbing states. For example, the states C and D can become absorbing if $p_{CC} = p_{DD} = 1$, in which case the invariant distribution $\boldsymbol{\pi}$ is such that $\pi_1 = \cdots = \pi_4 = 0$ (provided there are no other absorbing states).

The invariant distribution gives insight into the project portfolio status that is approached after many transitions, since

$$\lim_{t\to\infty} \boldsymbol{q}^{(t)} = \boldsymbol{q}\,\mathbf{P}^t = \boldsymbol{\pi},$$

where $\boldsymbol{q} \in \Delta_n$ is the initial project portfolio status.

Example 6. The invariant distribution of $\mathbf{P}$ in Ex. 4 is

$$\boldsymbol{\pi} = (0.1527, 0.1841, 0.0508, 0.0223, 0.1737, 0.4162);$$

cf. footnote 5. This means that, for example, in the long run 15.27% of projects will be in the *New* state, 18.41% will be actively progressing in *Green*, and so on.

C. Error Propagation

Since the time-t distribution $\boldsymbol{q}^{(t)}$ in Eq. (1), which tends towards the invariant distribution $\boldsymbol{\pi}$, depends less and less on the initial distribution $\boldsymbol{q}$, estimation errors in the initial distribution tend to taper out over time, contrary to the standard intuition of error propagation in systems. To see this, consider a decomposition of the transition matrix, in the form $\mathbf{P} = \mathbf{V}^{-1}\boldsymbol{\Lambda}\mathbf{V}$, where $\boldsymbol{\Lambda} = \mathrm{diag}(1, \lambda_2, \ldots, \lambda_n)$ is a diagonal matrix containing all eigenvalues of $\mathbf{P}$,[7] and where

$$\mathbf{V} = \begin{bmatrix} \boldsymbol{\pi} \\ \boldsymbol{v}_2 \\ \cdots \\ \boldsymbol{v}_n \end{bmatrix} \in \mathbb{C}^{n\times n},$$

with $\boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ the (possibly complex-valued) eigenvectors of $\mathbf{P}$ (in addition to $\boldsymbol{\pi}$). Then for any $\boldsymbol{q} \in \Delta_n$:

$$\boldsymbol{q}^{(t)} = \boldsymbol{q}\mathbf{P}^t = \boldsymbol{q}\mathbf{V}^{-1}\boldsymbol{\Lambda}^t\mathbf{V} = \boldsymbol{\pi} + \sum_{i=2}^{n} c_i \lambda_i^t \boldsymbol{v}_i, \quad t \geq 0,$$

where $(1, c_2, \ldots, c_n) = \boldsymbol{q}\mathbf{V}^{-1}$. Thus, considering an erroneous estimate $\hat{\boldsymbol{q}}$ of the initial distribution $\boldsymbol{q}$, their difference evolves according to

$$\boldsymbol{q}^{(t)} - \hat{\boldsymbol{q}}^{(t)} = \sum_{i=2}^{n} (c_i - \hat{c}_i)\lambda_i^t \boldsymbol{v}_i, \quad t \geq 0,$$

where $(1, \hat{c}_2, \ldots, \hat{c}_n) = \hat{\boldsymbol{q}}\mathbf{V}^{-1}$. More importantly,

$$\|\boldsymbol{q} - \hat{\boldsymbol{q}}\|_\infty \leq \varepsilon \;\Rightarrow\; \|\boldsymbol{q}^{(t)} - \hat{\boldsymbol{q}}^{(t)}\|_\infty \leq \varepsilon \max_{i\in\{2,\ldots,n\}} |\lambda_i|^t < \varepsilon,$$

for any $\boldsymbol{q}, \hat{\boldsymbol{q}} \in \Delta_n$, $t \in \{1, 2, \ldots\}$, and $\varepsilon \in (0, 1)$. This means that errors in the project portfolio state decrease exponentially.

Example 7. For $\hat{\boldsymbol{q}} = (25\%, 25\%, 5\%, 10\%, 17\%, 18\%) \in \Delta_n$ and $\boldsymbol{q}$ as in Ex. 5, it is $\varepsilon = \|\boldsymbol{q} - \hat{\boldsymbol{q}}\|_\infty = 5\%$. Under $\mathbf{P}$ given in Ex. 4, the maximum error in a 6-month-ahead forecast based on $\hat{\boldsymbol{q}}$ instead of $\boldsymbol{q}$ is $\|(\boldsymbol{q} - \hat{\boldsymbol{q}})\mathbf{P}^6\|_\infty \approx 1.53\% \leq \varepsilon\ell^6 \approx 3.46\%$, where $\ell = \max_{i\in\{2,\ldots,6\}} |\lambda_i| \approx 0.9403$. Thus, an identification error of 5% in each entry of the initial portfolio distribution provides a better-than-3.5% error guarantee in the entries of the 6-month forecast distribution.

D. Backcasting

Assume that the invariant distribution has full support (i.e., $\boldsymbol{\pi} \gg 0$) and that the current time-zero distribution is $\boldsymbol{\pi}$. If the current state is s_i, then the probability that at time $-t$ the project was in state s_j is

$$r_{ij}^{(t)} = \frac{\pi_j p_{ji}^{(t)}}{\pi_i}, \quad i, j \in \mathcal{N}, \quad t \in \{1, 2, \ldots\}.$$

If we set $\mathbf{R} = [r_{ij}]$ with $r_{ij} = r_{ij}^{(1)}$, then the backward analysis corresponds exactly to the forward analysis. As before,

$$\lim_{t\to\infty} r_{ij}^{(t)} = \frac{\pi_j \pi_i}{\pi_i} = \pi_j, \quad i, j \in \mathcal{N}.$$

Let $g_{jk}^{(t)}$ be the probability that when currently in state s_j a project came from state s_k without recurrence in exactly $t \geq 1$ steps. Then, analogous to the forward analysis in Sec. III-A,

$$g_{jk}^{(t)} = r_{jk}^{(t)} - \sum_{\tau=1}^{t-1} g_{jk}^{(\tau)} p_{kk}^{(t-\tau)}, \quad t \in \{1, 2, \ldots\}, \tag{4}$$

[5] Eqs. (2) and (3) yield the invariant distribution $\boldsymbol{\pi} = (0, \ldots, 0, 1)\,\boldsymbol{\Pi}^{-1}$, where

$$\boldsymbol{\Pi} = \begin{bmatrix} p_{11}-1 & p_{12} & \cdots & \cdots & p_{1,n-1} & 1 \\ p_{21} & p_{22}-1 & p_{23} & \cdots & p_{2,n-1} & 1 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ p_{n1} & \cdots & \cdots & \cdots & p_{n,n-1} & 1 \end{bmatrix}$$

is obtained by replacing the last column in $\mathbf{P} - \mathbf{I}$ with ones, where $\mathbf{I}$ denotes the $(n \times n)$-identity matrix. Indeed, $\boldsymbol{\pi}$ corresponds to the last line of $\boldsymbol{\Pi}^{-1}$.

[6] The transition matrix $\mathbf{P}$ is *aperiodic* if there is no $\tau > 0$ and no $i \in \mathcal{N}$ such that $p_{ii}^{(t)} = 0 \Leftrightarrow t \neq \nu\tau$, $\nu \in \{1, 2, \ldots\}$; cf. [11, p. 387].

[7] Exactly one eigenvalue of $\mathbf{P}$ is equal to 1, by Eq. (2), and the other $n-1$ eigenvalues $\lambda_2, \ldots, \lambda_n \in \mathbb{C}$ are such that $\max_{i\in\{2,\ldots,n\}} |\lambda_i| < 1$.

so, for example, $g_{jk}^{(1)} = r_{jk}^{(1)} = (\pi_k/\pi_j)\, p_{kj}$, for all $j, k \in \mathcal{N}$.

IV. Portfolio Performance

The analysis performed thus far to characterize the evolution of a PPMC can be used to determine important KPIs of the firm's project portfolio, such as success rate, project duration, failure time, the effectiveness of flagging projects, expected transition times, and the overall project-resource use both in terms of the cyclicality of project resources and the balance of a random project unit compared to its line duties elsewhere in the organization.

A. Long-Term Success Rate

The (long-term) success rate for projects is arguably different from the project-completion rate (π_{n-1}) of the steady-state distribution $\boldsymbol{\pi}$ introduced in Sec. III-B. The reason is that at any given time the transitory (or stationary) project portfolio may contain numerous ongoing projects (with states in $\mathcal{S}$) or projects in the planning phase (i.e., in the service state s_N). Hence, it is necessary to disable the resource feedback loop in $\mathbf{P}$ by setting $(p_{CN}, p_{CC}) = (p_{DN}, p_{DD}) = (0, 1)$. This converts the service states *Completed* and *Discontinued* into absorbing states; cf. footnote 3. Computing the limiting distribution,

$$\hat{\boldsymbol{\pi}} = (\hat{\pi}_1, \ldots, \hat{\pi}_n) = \lim_{t\to\infty} \boldsymbol{q}\, \hat{\mathbf{P}}^t,$$

for the initial distribution $\boldsymbol{q} = (1, 0, \ldots, 0)$ yields the *success rate* $\hat{\pi}_C$ (together with the *failure rate* $\hat{\pi}_D = 1 - \hat{\pi}_C$), where $\hat{\mathbf{P}}$ is the modified transition matrix which contains the aforementioned adjustments to the last two rows of $\mathbf{P}$.

Example 8. Based on $\mathbf{P}$ in Ex. 4, converting s_C and s_D into absorbing states leads to the modified transition matrix,

$$\hat{\mathbf{P}} = \begin{bmatrix} 0.75 & 0.15 & 0 & 0 & 0 & 0.10 \\ 0 & 0.79 & 0.10 & 0.02 & 0.08 & 0.01 \\ 0 & 0.20 & 0.55 & 0.20 & 0.03 & 0.02 \\ 0 & 0.25 & 0.20 & 0.38 & 0.05 & 0.12 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix},$$

with limiting distribution $\hat{\boldsymbol{\pi}} \approx (0, \ldots, 0, 45.49\%, 54.51\%)$, implying an overall success rate $\hat{\pi}_C$ of just over 45%.[8]

B. Project Duration

The *project duration* τ_C *of successful* (i.e., eventually completed) *projects* is the expected time it takes to go from state s_G to state s_C for the first time; similarly, the *project duration of unsuccessful* (i.e., discontinued) *projects* is the expected time to go from s_G to s_D for the first time, without passing via any other service state. That is using the same modified transition matrix $\hat{\mathbf{P}}$ as in Sec. IV-A with the corresponding first-passage distributions as introduced in Sec. III-A, it is $\tau_C = \mu_{GC}$ and $\tau_D = \mu_{GN}$, where

$$\mu_{jk} = \sum_{t=1}^{\infty} t\, (f_{jk}^{(t)}/f_{jk}), \quad j, k \in \mathcal{N}, \tag{5}$$

is the mean transition time between two states s_j and s_k, conditionally on eventually arriving in state s_k.[9] Hence, the overall project duration (no matter if unsuccessful or not) is

$$\bar{\tau} = \sum_{t=1}^{\infty} t \left(f_{GC}^{(t)} + f_{GD}^{(t)} \right) = f_{GC}\, \tau_C + (1 - f_{GC})\, \tau_D,$$

amounting to a weighted average of the conditional project durations.

Example 9. For the modified transition matrix $\hat{\mathbf{P}}$ (without resource feedback) in Ex. 8, we find $\tau_C = \mu_{GC} \approx 10.8$ periods and $\tau_D = \mu_{GD} \approx 12.7$ periods, meaning that discontinued projects linger in the organization for just over 12 months, longer than successfully completed projects ($\bar{\tau} \approx 11.2$ periods). Naturally, $f_{GC} = \hat{\pi}_C \approx 75.82\%$ and $f_{GD} = \hat{\pi}_D \approx 24.18\%$ (cf. footnote 8), so $f_{GC} + f_{GD} = 1$.

C. Failure Time

The average time to failure from the initiation of the project corresponds to $\tau_D = \mu_{GD}$ as in Eq. (5). Conditional on having been flagged as yellow, the time to failure becomes μ_{YD}, and conditional on a red flag, it is μ_{GD}. The corresponding calculations should be performed using the modified transition matrix $\hat{\mathbf{P}}$ discussed in Sec. IV-A.

Example 10. Given $\hat{\mathbf{P}}$ as in Ex. 8, one obtains $\mu_{YD} \approx 10.4$ periods for the expected failure time from the state s_Y, and $\mu_{RD} \approx 7.5$ periods for the expected failure time from the state s_R. The average duration of unsuccessful projects τ_D for this transition matrix was computed in Ex. 9.

D. Flag Effectiveness

To determine the effectiveness of the traffic-light approach it is useful to assess whether the flagged project states s_Y and s_R serve as effective warnings over a given time horizon T. For this consider the random state $\tilde{s}_G^{(t)}$ of a project t periods after having been started in *Green*, so $\tilde{s}_G^{(0)} = s_G$, under the PPMC $\hat{\mathbf{P}}$ as in Sec. IV-A (obtained from $\mathbf{P}$ by disabling resource feedback). The (*T-period*) *sensitivity* of flag $i \in \{Y, R\}$,

$$\theta_i^{(T)} = \mathbb{P}(s_i \in \{\tilde{s}_G^{(1)}, \ldots, \tilde{s}_G^{(T-1)}\} | \tilde{s}_G^{(T)} = s_D),$$

is the probability of the flag's appearance, conditional on the project having reached the absorbing state s_D in T periods or less. Conversely, the (*T-period*) *specificity* of flag $i \in \{Y, R\}$,

$$\vartheta_i^{(T)} = \mathbb{P}(s_i \notin \{\tilde{s}_G^{(1)}, \ldots, \tilde{s}_G^{(T-1)}\} | \tilde{s}_G^{(T)} = s_C),$$

is the probability of flag i's absence, conditional on the project reaching successful completion (s_C) after T periods or less.

[8] This success rate is biased downward by the fact that some projects get discontinued at the planning stage. To eliminate this effect, one can consider the limiting distribution for projects conditionally on having been launched (in the state s_G), $\lim_{t\to\infty}(0, 1, 0, 0, 0, 0)\, \hat{\mathbf{P}}^t \approx (0, 0, 0, 0, 75.82\%, 24.18\%)$, which then yields a much higher success rate, in excess of 75%.

[9] Since the modified transition matrix $\hat{\mathbf{P}}$ is not irreducible, it is generically $f_{GC} = 1 - f_{GD} = \sum_{t=0}^{\infty} f_{GC}^{(t)} < 1$.

Using the first-passage distribution $\hat{f}_{ij}^{(t)}$ with respect to $\hat{\mathbf{P}}$ (as introduced in Sec. III-A with respect to $\mathbf{P}$), we obtain that

$$\theta_i^{(T)} = \frac{\sum_{t=1}^{T-1} \hat{f}_{Gi}^{(t)} \sum_{\tau=1}^{T-t} \hat{f}_{iD}^{(\tau)}}{\sum_{t=1}^{T} \hat{f}_{GD}^{(t)}}, \quad T \geq 1,$$

and

$$\vartheta_i^{(T)} = \frac{\sum_{t=1}^{T} \hat{f}_{GC}^{(t)} - \sum_{t=1}^{T-1} \hat{f}_{Gi}^{(t)} \sum_{\tau=1}^{T-t} \hat{f}_{iC}^{(\tau)}}{\sum_{t=1}^{T} \hat{f}_{GC}^{(t)}}, \quad T \geq 1,$$

for flag $i \in \{Y, R\}$.

Example 11. Given the modified transition matrix $\hat{\mathbf{P}}$ in Ex. 8, the T-period (sensitivity, specificity)-tuples for the flag states s_Y and s_R are given in Table I.

TABLE I
T-PERIOD FLAG-EFFECTIVENESS INDICATORS FOR s_Y AND s_R.

T	$(\theta_Y^{(T)}, \vartheta_Y^{(T)})$	$(\theta_R^{(T)}, \vartheta_R^{(T)})$
6	(41.2%, 86.2%)	(38.6%, 93.8%)
12	(57.8%, 71.8%)	(51.5%, 85.1%)
18	(64.5%, 63.5%)	(57.0%, 78.9%)
∞	(71.6%, 53.3%)	(63.9%, 69.0%)

This means that for the given PPMC the flagged project state *Yellow* exhibits a higher sensitivity paired with a lower specificity than the flagged project state *Red*. The quality of either flag as a predictor of medium-term project failure (say, with a year) is only moderate.[10]

E. Resource Cycles

The average duration of a successful project is $\tau_C = \mu_{GC}$ as noted in Sec. IV-B. Its resources then spend μ_{CG} periods outside projects.[11] Conversely, the average duration of a failed project (conditional on having been initialized) is $\tau_D = \mu_{GD}$, with its resources being redeployed elsewhere for an average time of μ_{DG}. Thus, the average cycle time is

$$\bar{\tau}_{\text{cycle}} = \hat{\pi}_C \left(\tau_C + \mu_{CG}\right) + \left(1 - \hat{\pi}_C\right)\left(\tau_D + \mu_{DG}\right), \quad (6)$$

by virtue of the ergodic theorem [16, Thm. 7].

Example 12. For $\mathbf{P}$ as in Ex. 4, Eq. (5) yields that $\mu_{CG} \approx 30$ periods and $\mu_{DG} \approx 32.8$ periods. Thus, the average cycle time in Eq. (6) becomes $\bar{\tau}_{\text{cycle}} \approx 41.9$ periods, where τ_C, τ_D are given in Ex. 9.

F. Resource Balance

Assuming no systematic effects of project size on its outcome, the ratio ρ of in-project resource use versus outside-project resource use can be based on the partial cycle times computed in Sec. IV-E, so

$$\rho = \frac{\hat{\pi}_C \tau_C + (1 - \hat{\pi}_C)\tau_D}{\hat{\pi}_C \left(\tau_C + \mu_{CG}\right) + \left(1 - \hat{\pi}_C\right)\left(\tau_D + \mu_{DG}\right)} = \frac{\bar{\tau}}{\bar{\tau}_{\text{cycle}}}.$$

[10] A project in state s_Y has after all a 67.2% success rate, and a project in state s_R succeeds with 60.7% probability (using computations as in footnote 8 for starting distributions that assign probability 1 to either s_Y or s_R).

[11] We abstract from the fact that project planning may involve project resources, although of course in some organizations this task is performed by separate entities, which is the view we take here, for simplicity.

Example 13. In continuation of Ex. 12, we find $\rho = \bar{\tau}/\bar{\tau}_{\text{cycle}} \approx 26.8\%$, where $\bar{\tau}$ was provided in Ex. 9 and $\bar{\tau}_{\text{cycle}}$ in Ex. 12. Hence, relative to this project portfolio, resources are used a bit more than a quarter of the time in projects and about three-quarters of the time outside projects (including project planning as part of these outside activities).

V. Conclusion

The Markov chain-based approach presented in this paper offers a structured, quantitative method for tracking and forecasting project portfolio status. By producing finite-time predictions and long-run state distributions, it enables decision-makers to anticipate trends, manage risks, and optimize resource allocation. Forecasts are shown to be reasonably robust to moderate errors in initial state assessments, strengthening the methods practical value in real-world settings. A limitation of this approach is its reliance on comparable project types, for example, by assuming similar dynamics across IT and non-IT projects. If such differences exist, the portfolio should be divided into *sub*portfolios of homogeneous projects. To ensure reliable aggregation, the classification of ongoing projects into transient states (in $\mathcal{S}$) must also be applied consistently across the organization. Beyond risk mitigation, this supports clearer, more effective communication about project portfolio health.

References

[1] H. L. Gantt, Work, Wages, and Profits, New York: Engineering Magazine, 1913.
[2] E. R. Marsh, "The harmonogram of Karol Adamiecki," Academy of Management Journal, vol. 18, no. 2, pp. 358–364, June 1975.
[3] H. M. Markowitz, "Portfolio selection," Journal of Finance, vol. 7, no. 1, pp. 77–91, March 1952.
[4] H. M. Markowitz, "Foundations of portfolio theory," Journal of Finance, vol. 46, no. 2, pp. 469–477, June 1991.
[5] R. G. Cooper, "Stage-gate systems: A new tool for managing new products," Business Horizons, Vol. 33, No. 3, pp. 44–54, May-June 1990.
[6] R. G. Cooper, "Perspective: The stage-gate® idea-to-launch process—Update, what's new, and NexGen systems," Journal of Product Innovation Management, vol. 25, no. 3, pp. 213–232, May 2008.
[7] E. C. Conforto, F. Salum, D. C. Amaral, S. L. Da Silva, and L. F. M. De Almeida, "Can agile project management be adopted by industries other than software development?" Project Management Journal, vol. 45, no. 3, pp. 21–34, 2014.
[8] R. P. Hardingham, "A simple model approach to multi-project monitoring," R&D Management, vol. l, no. l, pp. 43–47, 1970.
[9] L. K. Hansen and P. Svejvig, "Seven decades of project portfolio management research (1950-2019) and Perspectives for the Future," Project Management Journal, vol. 53, no. 3, pp. 277–294, April 2022.
[10] M. Jeffery and I. Leliveld, "Best practices in IT portfolio management," MIT Sloan Management Review, vol. 45, Spring Issue, pp. 41–49, April 2004.
[11] W. Feller, An Introduction to Probability Theory and Its Applications, Vol. I (Third Ed.), New York: Wiley, 1968.
[12] V. G. Kulkarni and V. G. Adlakha, "Markov and Markov-regenerative PERT networks," Operations Research, vol. 34, no. 5, pp. 769–781, September-October 1986.
[13] B. Hermans and R. Leus, "Scheduling Markovian PERT networks to maximize the net present value: New results," Operations Research Letters, vol. 46, no. 2, pp. 240–244, March 2018.
[14] US Navy, "Program evaluation research task," Summary Report, Special Projects Office, Bureau of Naval Weapons, Washington, DC, July 1958.
[15] J. G. Kemeny and J. L. Snell, Finite Markov Chains, Princeton, NJ: Van Nostrand, 1960.
[16] W. L. Smith, "Regenerative Stochastic Processes," Proceedings of the Royal Society of London (Series A), vol. 232, no. 1188, pp. 6–31, October 1955.