[Supplementary Material] Disease gene prioritization by integrating tissue-specific molecular networks using a robust multi-network model

Jingchao Ni¹, Mehmet Koyuturk¹, Hanghang Tong², Jonathan Haines³, Rong Xu³ and Xiang Zhang⁴*

Correspondence: xzhang@ist.psu.edu

¹ College of Information Sciences and Technology, Pennsylvania State University, 332 Information Sciences and Technology Building, PA 16802, University Park, USA

Summary
In this supplementary material, we provide the matrix form of \( J_{CR} \), the optimization solution to \( J_{CR} \), the algorithm and the theoretical analysis of CR.

Algorithm CR

| Algorithm 1: CR |
|----------------|
| **Input:** (1) a disease similarity network \( A \); (2) the tissue-specific molecular networks \( \{ G_i \} \); (3) the seed vectors \( \{ e_i \} \); and (4) the parameters \( \beta \) and \( c \)  
| **Output:** the ranking vectors \( r_1, \ldots, r_h \) |
| **Offline-computation:** Construct \( \tilde{G} \) and \( \tilde{Y} \) from \( A \) and \( \{ G_i \} \); |
| **Online-ranking:** |
| 3 Construct the aggregated seed vector \( e = (e_1^T, \ldots, e_h^T)^T \); |
| 4 Initialize the aggregated ranking vector \( r = e \); |
| 5 while not convergence do |
| 6 Update: \( r \leftarrow (c + 2\beta)\tilde{G} + \frac{2\beta}{1 + 2\beta}\tilde{Y})r + \frac{1 - c}{1 + 2\beta}e; \) |
| 7 end |
| 8 return the ranking vectors \( r_1, \ldots, r_h \) based on \( r \) |

Complexity Analysis of CR

Let \( n_i \) be the number of nodes in \( G_i \) and \( n = \sum_{i=1}^{h} n_i \). Let \( m_i \) be the number of edges in \( G_i \) and \( m = \sum_{i=1}^{h} m_i \). There are \( O(m + hn) \) nonzero entries in \( \tilde{G} \) and \( \tilde{Y} \) in total. Thus the offline-computation and online-ranking time complexities of Algorithm 1 are \( O(m + hn) \) and \( O(T^*(m + hn)) \), respectively, where \( T^* \) is the total number of iterations. Since we need to store \( \tilde{G} \) and \( \tilde{Y} \), thus the space complexity is \( O(m + hn) \).

Generally, \( h \) is much smaller than \( n \) and can be regarded as constants. Hence we can regard the time and space complexities of Algorithm 1 as \( O(T^*(m + n)) \) and \( O(m + n) \), respectively.

Matrix Form of \( J_{CR} \)

The objective function \( J_{CR} \) is jointly convex in \( r_1, \ldots, r_h \). This can be shown by first deriving its matrix form.
Let \( r = (r_1^T, ..., r_h^T)^T \), \( e = (e_1^T, ..., e_h^T)^T \), i.e., we concatenate all ranking and seed vectors. Let \( \tilde{G} = \text{diag}(\tilde{G}_1, ..., \tilde{G}_h) \) be a diagonal block matrix. Then we have

\[
\text{cr}^T(I_n - \tilde{G})r + (1 - c)\|r - e\|_F^2 = \sum_{i=1}^h \Theta_{\text{within}}(r_i) \tag{1}
\]

where \( I_n \) is an \( n \times n \) identity matrix and \( n = \sum_{i=1}^h n_i \).

Define a common gene mapping matrix \( O_{ij} \in \{0, 1\}^{n_i \times n_j} \) where \( O_{ij}(x, y) = 1 \) if node \( x \) in \( G_i \) and node \( y \) in \( G_j \) represent the same gene; \( O_{ij} = 0 \) otherwise. Then \( Y \) is a block matrix whose \((i, j)^{th}\) block is \( A(i, j)O_{ij} \). Note that \( A(i, i) = 0 \).

Further, let \( D_V = \text{diag}(d_A(1)I_{n_1}, ..., d_A(h)I_{n_h}) \) be a diagonal matrix, where \( d_A(i) = \sum_{j=1}^h A(i, j) \). We define \( X = D_V^{-\frac{1}{2}}(D_V - Y)D_V^{-\frac{1}{2}} = I_n - \tilde{Y} \), where \( \tilde{Y} = D_V^{-\frac{1}{2}}YD_V^{-\frac{1}{2}} \).

We have

\[
r^T Xr = \frac{1}{2} \sum_{i,j=1}^h \Theta_{\text{cross}}(r_i, r_j) \tag{2}
\]

According to Eq. (1) and Eq. (2), we have the following theorem.

**Theorem 1 Matrix Form of \( J_{CR} \).** \( J_{CR} \) has the following matrix form

\[
\min_{r \geq 0} J_{CR} = \text{cr}^T(I_n - \tilde{G})r + (1 - c)\|r - e\|_F^2 + 2\beta r^T Xr \tag{3}
\]

**Proof** The proof of Theorem 1 includes two equivalence validations:

1. \( \text{cr}^T(I_n - \tilde{G})r + (1 - c)\|r - e\|_F^2 = \sum_{i=1}^h \Theta_{\text{within}}(r_i) \)
2. \( r^T Xr = \frac{1}{2} \sum_{i,j=1}^h \Theta_{\text{cross}}(r_i, r_j) \)

Since the equivalence (1) is obvious, we only need to prove the equivalence (2).

According to the definition of \( X \) and \( r \), we have

\[
r^T Xr = r^T I_n r - r^T \tilde{Y}r = \sum_{i=1}^h r_i^T I_{n_i} r_i - \sum_{i,j=1}^h r_i^T \tilde{Y}_{ij} r_j \tag{4}
\]

where \( \tilde{Y}_{ij} \in \mathbb{R}_{+}^{n_i \times n_j} \) is the \((i, j)^{th}\) block of \( \tilde{Y} \). Note \( \tilde{Y}_{ii} = 0 \) \((1 \leq i \leq h)\). Then let \( (D_V)_i \) be the \( i^{th} \) diagonal block of \( D_V \) and \( Y_{ij} \) be the \((i, j)^{th}\) block of \( Y \). Recall \( Y_{ij} = A(i, j)O_{ij} \). We have

\[
r^T Xr = \frac{1}{2} \left( \sum_{i=1}^{h} \frac{r_i^T}{\sqrt{d_A(i)}} (D_V)_i \cdot \frac{r_i}{\sqrt{d_A(i)}} - 2 \sum_{i,j=1}^{h} \frac{r_i^T}{\sqrt{d_A(i)}} Y_{ij} \cdot \frac{r_j}{\sqrt{d_A(j)}} + \sum_{j=1}^{h} \frac{r_j^T}{\sqrt{d_A(j)}} (D_V)_j \cdot \frac{r_j}{\sqrt{d_A(j)}} \right)
\]

Let \( D_Y \) be the degree matrix of \( Y \) and \( (D_Y)_i \) be the \( i^{th} \) diagonal block of \( D_Y \).

Define \( D_{Y_{ij}} \) to be the degree matrix of \( Y_{ij} \) (note the nonzero diagonal values of \( D_{Y_{ij}} \) are \( A(i, j) \)). Then \( (D_Y)_i = \sum_{j=1}^{h} D_{Y_{ij}} \). Define \( D_{Y_{ij}} \) to be an \( n_i \times n_i \) diagonal matrix s.t. \( D_{Y_{ij}} + D_{Y_{ij}} = A(i, j)I_{n_i} \). Then let \( (D_Y)_i = \sum_{j=1}^{h} D_{Y_{ij}} \), we
have \((D_Y)_i = (D_Y + D_Y)_i\). Thus

\[
\begin{aligned}
    r^T X r &= \frac{1}{2} \sum_{i=1}^{h} \frac{r_i^T}{\sqrt{d_A(i)}} (D_Y + D_Y)_i \frac{r_i}{\sqrt{d_A(i)}} - 2 \sum_{i,j=1}^{h} \frac{r_i^T}{\sqrt{d_A(i)}} Y_{ij} \frac{r_j}{\sqrt{d_A(j)}} \\
    &+ \sum_{j=1}^{h} \frac{r_j^T}{\sqrt{d_A(j)}} (D_Y + D_Y)_j \frac{r_j}{\sqrt{d_A(j)}} \\
    &= \frac{1}{2} \sum_{i,j=1}^{h} A(i,j) \left( \frac{r_i^T (I_{ij})}{\sqrt{d_A(i)}} \frac{r_i (I_{ij})}{\sqrt{d_A(i)}} - 2 \frac{r_i^T (I_{ij})}{\sqrt{d_A(i)}} \frac{r_j (I_{ij})}{\sqrt{d_A(j)}} + \frac{r_j^T (I_{ij})}{\sqrt{d_A(j)}} \frac{r_j (I_{ij})}{\sqrt{d_A(j)}} \right) \\
    &+ \frac{\sum_{j=1}^{h} A(i,j) (\frac{r_i (I_{ij})}{\sqrt{d_A(i)}} \frac{r_j (I_{ij})}{\sqrt{d_A(j)}} + \frac{r_j (I_{ij})}{\sqrt{d_A(j)}} \frac{r_j (I_{ij})}{\sqrt{d_A(j)}})}{\sqrt{d_A(i)}} \\
    &= \frac{1}{2} \sum_{i,j=1}^{h} A(i,j) \left( \| \frac{r_i (I_{ij})}{\sqrt{d_A(i)}} \frac{r_j (I_{ij})}{\sqrt{d_A(j)}} \|_F^2 + \| \frac{r_j (I_{ij})}{\sqrt{d_A(j)}} \|_F^2 + \| \frac{r_j (I_{ij})}{\sqrt{d_A(j)}} \|_F^2 \right) \\
    &= \frac{1}{2} \sum_{i,j=1}^{h} \Theta_{cross}(r_i, r_j)
\end{aligned}
\]

This completes the proof. \(\square\)

**Optimization Solution to J_{CR}**

From Theorem 1, \(J_{CR}\) is a quadratic function of \(r\). We can derive a power method to minimize \(J_{CR}\) as follows.

\[
\frac{\partial J_{CR}}{\partial r} = 2 \left( (1 + 2\beta) I_n - (c\tilde{G} + 2\beta \tilde{Y}) \right) r - 2(1 - c)e
\]

Using gradient descent, if we set \(r \leftarrow r - \eta \frac{\partial J_{CR}}{\partial r}\), where \(\eta = \frac{1}{2(1 + 2\beta)}\), we have

\[
r \leftarrow \left( \frac{c}{1 + 2\beta} \tilde{G} + \frac{2\beta}{1 + 2\beta} \tilde{Y} \right) r + \frac{1 - c}{1 + 2\beta} e
\]  \(\text{(5)}\)

Eq. (5) is a fixed-point approach to compute \(r\) that converges to the global optimal solution of \(J_{CR}\). Algorithm 1 summarizes our approach according to the optimization solution.

**Theoretical Analysis of CR**

In this section, we show that Algorithm 1 converges to the global minimum of \(J_{CR}\) by Theorem 2 and Theorem 3.

**Theorem 2** **Convergence of CR.** Algorithm 1 converges to the closed-form solution

\[
r = (I_n - \frac{c}{1 + 2\beta} \tilde{G} - \frac{2\beta}{1 + 2\beta} \tilde{Y})^{-1} \frac{1 - c}{1 + 2\beta} e
\]
Proof First, the closed-form solution can be obtained by solving $\frac{\partial \ln \text{tr}}{\partial r} = 0$. Then let $\mathbf{M} = \frac{c}{1 + 2\beta} \tilde{\mathbf{G}} + \frac{2\beta}{1 + 2\beta} \tilde{\mathbf{Y}}$, the CR updating rule in Eq. (5) becomes $\mathbf{r} = \mathbf{M} \mathbf{r} + \frac{1 - c}{1 + 2\beta} \mathbf{e}$. Next, we show that the eigenvalues of $\mathbf{M}$ are in the range of $(-1, 1)$.

Let $\mathbf{G} = \text{diag}(\mathbf{G}_1, ..., \mathbf{G}_h)$ and $\mathbf{D}_G$ be its degree matrix, then $\tilde{\mathbf{G}} = \mathbf{D}_G^{-\frac{1}{2}} \mathbf{GD}_G^{-\frac{1}{2}}$. Since $\tilde{\mathbf{G}}$ is similar to the stochastic matrix $\mathbf{G}\mathbf{D}_G^{-1} = \mathbf{D}_G^{-\frac{1}{2}} \tilde{\mathbf{G}} \mathbf{D}_G^{-\frac{1}{2}}$, it has eigenvalues within $[-1, 1]$. Also, $\tilde{\mathbf{Y}}$ is similar to the matrix $\mathbf{YD}_V^{-1} = \mathbf{D}_V^{-\frac{1}{2}} \tilde{\mathbf{Y}} \mathbf{D}_V^{-\frac{1}{2}}$ where each column sum of $\mathbf{YD}_V^{-1}$ is no greater than 1.

The Gershgorin Circle Theorem [1] states that for a complex $n \times n$ matrix $\mathbf{B}$, every eigenvalue $\lambda$ of $\mathbf{B}$ lies within at least one of the Gershgorin discs $\{ \lambda : |\lambda - b_{ii}| \leq \sum_{j=1, j \neq i}^{n} |b_{ji}| \} (i = 1, ..., n)$, where $b_{ii}$ is the $i$th diagonal value of $\mathbf{B}$ and $b_{ji}$ is the $(j, i)$th entry of $\mathbf{B}$. Since $\mathbf{A}(i, i) = 0$ for $i = 1, ..., h$, the diagonal values of $\mathbf{Y}$ are zero. Therefore, the eigenvalues of $\mathbf{YD}_V^{-1}$ satisfy $|\lambda| \leq 1$, which implies the eigenvalues of $\tilde{\mathbf{Y}}$ are within $[-1, 1]$.

One result of the Weyl’s Inequality Theorem [2] states that for matrices $\tilde{\mathbf{H}}, \mathbf{H}, \mathbf{P} \in \mathcal{H}_n$, where $\mathcal{H}_n$ is the set of $n \times n$ Hermitian matrices, if $\tilde{\mathbf{H}} = \mathbf{H} + \mathbf{P}$ and their eigenvalues are arranged in non-increasing orders, i.e., $\lambda_1(\tilde{\mathbf{H}}) \geq ... \geq \lambda_n(\tilde{\mathbf{H}})$, $\lambda_1(\mathbf{H}) \geq ... \geq \lambda_n(\mathbf{H})$, $\lambda_1(\mathbf{P}) \geq ... \geq \lambda_n(\mathbf{P})$, then the following inequalities hold:

$$\lambda_n(\mathbf{P}) \leq \lambda_i(\tilde{\mathbf{H}}) - \lambda_i(\mathbf{H}) \leq \lambda_i(\mathbf{P}), \forall i = 1, ..., n$$

Since $\tilde{\mathbf{G}}, \tilde{\mathbf{Y}}, \mathbf{M} \in \mathcal{H}_n$ and $\mathbf{M} = \frac{c}{1 + 2\beta} \tilde{\mathbf{G}} + \frac{2\beta}{1 + 2\beta} \tilde{\mathbf{Y}}$, we have

$$\lambda_1(\mathbf{M}) \leq \frac{c}{1 + 2\beta} \lambda_1(\tilde{\mathbf{G}}) + \frac{2\beta}{1 + 2\beta} \lambda_1(\tilde{\mathbf{Y}})$$

$$\lambda_n(\mathbf{M}) \geq \frac{c}{1 + 2\beta} \lambda_n(\tilde{\mathbf{G}}) + \frac{2\beta}{1 + 2\beta} \lambda_n(\tilde{\mathbf{Y}})$$

which means the eigenvalues of $\mathbf{M}$ are in the range of $[-\frac{c + 2\beta}{1 + 2\beta}, \frac{c + 2\beta}{1 + 2\beta}]$. Since $0 < c < 1$, the eigenvalues of $\mathbf{M}$ are in the range of $(-1, 1)$.

Based on this property, we can show the convergence of the fixed-point approach. Without loss of generality, let $\mathbf{r}^{(0)} = \mathbf{e}$, and $t$ be the iteration index ($t \geq 1$). According to the CR updating rule in Eq. (5), we have

$$\mathbf{r}^{(t)} = \mathbf{M}^{t} \mathbf{e} + \sum_{i=0}^{t-1} \mathbf{M}^i \frac{1 - c}{1 + 2\beta} \mathbf{e}$$

Since the eigenvalues of $\mathbf{M}$ are all in $(-1, 1)$, we have

$$\lim_{t \to \infty} \mathbf{M}^t = 0, \text{ and } \lim_{t \to \infty} \sum_{i=0}^{t-1} \mathbf{M}^i = (\mathbf{I}_n - \mathbf{M})^{-1}$$

Therefore

$$\mathbf{r} = \lim_{t \to \infty} \mathbf{r}^{(t)} = (\mathbf{I}_n - \mathbf{M})^{-1} \frac{1 - c}{1 + 2\beta} \mathbf{e} = (\mathbf{I}_n - \frac{c}{1 + 2\beta} \tilde{\mathbf{G}} - \frac{2\beta}{1 + 2\beta} \tilde{\mathbf{Y}})^{-1} \frac{1 - c}{1 + 2\beta} \mathbf{e}$$

which is the closed-form solution. \qed
Theorem 3  Optimality of CR. At convergence, Algorithm 1 gives the global minimum of $J_{CR}$ defined in Eq. (3).

Proof This can be proved by showing that the function in Eq. (3) is convex. The Hessian matrix of Eq. (3) is $\nabla^2 J_{CR} = 2((1 + 2\beta)I_n - (c\tilde{G} + 2\beta\tilde{Y}))$. Following the similar idea as in the proof of Theorem 2, we have that the eigenvalues of $\nabla^2 J_{CR}$ are no less than $2(1 - c)$. Since $0 < c < 1$, $\nabla^2 J_{CR}$ is positive-definite. Therefore, Eq. (3) is convex.