A Proofs

A.1 Proof of Proposition 1

Proof. To show that the state transition equation (Formula 1) gives the minimum repair cost in each step, suppose that $dp(i,j)$ and its corresponding match M is not the minimum repair cost, i.e., there exists another match $M' \neq M$ that generates a lower cost $c' < dp(i,j)$. We will discuss the last element in M', i.e., the last match pair $(x', y') \in M$.

(1) If $(x', y') = (t_i, s_j)$, then we can remove the last element, thus having $c' - \Delta_m(t_i, s_j) < dp(i,j) - \Delta_m(t_{i-1}, s_{j-1})$. However, since $dp(i-1, j-1)$ is the minimum repair cost from $\mathbf{t}_{[i-1]}$ to $\mathbf{s}_{[j-1]}$, we have $dp(i-1,j-1) \leq c' - \Delta_m(t_{i-1}, s_{j-1}) < dp(i,j) - \Delta_m(t_{i-1}, s_{j-1})$, i.e., $dp(i-1,j-1) + \Delta_m(t_{i-1}, s_{j-1}) < dp(i,j)$. However, according to the condition in Formular 1, $dp(i,j) \leq dp(i-1,j-1) + \Delta_m(t_{i-1}, s_{j-1})$, thus contradicts.

(2) If $(x', y') = (t_i, *)$ or $(x', y') = (*, s_j)$, analogously, we could obtain $dp(i-1,j) + \Delta_d(t_{i-1}) < dp(i,j)$ or $dp(i, j-1) + \Delta_a(s_{i-1}) < dp(i,j)$ respectively, which again leads to the contradiction.

Therefore, we can conclude the correctness of the state transition equation $dp(i,j)$, which gives the minimum repair cost from time series $\mathbf{t}_{[i]}$ to $\mathbf{s}_{[j]}$. □

A.2 Proof of Lemma 2

Proof. If $m^* > \lfloor \frac{c}{\lambda_a} + n \rfloor$, we have $m^* - n > \lfloor \frac{c}{\lambda_a} \rfloor$ (since m^*, n are integers, it is equivalent to $m^* - n > \lfloor \frac{c}{\lambda_a} \rfloor$), i.e., $\mathbf{s}$ has at least $(\lfloor \frac{c}{\lambda_a} \rfloor + 1)$ more points than $\mathbf{t}$. To match the two sequences, at least $(\lfloor \frac{c}{\lambda_a} \rfloor + 1)$ times of insert operations should be conducted, which can generate $(\lfloor \frac{c}{\lambda_a} \rfloor + 1) * \lambda_a$ repair cost. Since $(\lfloor \frac{c}{\lambda_a} \rfloor + 1) * \lambda_a > c$, we find the result that there could not exist lower cost with $m^* > \lfloor \frac{c}{\lambda_a} + n \rfloor$, i.e., $m^* \leq \lfloor \frac{c}{\lambda_a} + n \rfloor$. □

A.3 Proof of Lemma 3

Proof. First, if s_0^* ¿ t_{n-1}, let m and M denote the length and the match of $\mathbf{s}$ that computed by Algorithm 1, respectively. Consider another regular interval time series $\mathbf{s}'$ of the same interval and length with $\mathbf{s}$, while $s'_0 = t_{n-1}$. That is to say, $s'_i \in \mathbf{s}'$ correspond one-to-one to $s_i \in \mathbf{s}$. We use M' to represent the match from $\mathbf{t}$ to $\mathbf{s}'$ that corresponds to M, i.e., replace the s_j in M with s'_j. According to Definition 8, since s_0^* ¿ t_{n-1}, we have $t_i < s_i$,

$$\begin{aligned}
\Delta(\mathbf{t}, \mathbf{s}', M') &= \sum_{(t_i, s'_j) \in M'} |t_i - s'_j| + \sum_{(t_i, *) \in M'} \lambda_d + \sum_{(*, s'_j) \in M'} \lambda_a \\
&= \sum_{(t_i, s'_j) \in M'} (t_i - s'_j) + \sum_{(t_i, *) \in M} \lambda_d + \sum_{(*, s_j) \in M} \lambda_a \\
&< \sum_{(t_i, s_j) \in M} (t_i - s_j) + \sum_{(t_i, *) \in M} \lambda_d + \sum_{(*, s_j) \in M} \lambda_a \\
&= \Delta(\mathbf{t}, \mathbf{s}, M)
\end{aligned}$$

Therefore, $\mathbf{s}'$ generates lower repair cost, which contradicts the assumption that $\mathbf{s}$ shows the minimum cost.

Next, if $s_0^* < t_0 - (\lambda_a + \lambda_d)$, it is easy to show that the subsequence $\mathbf{s}' = \{s_1, s_2, \ldots, s_{n-1}\}$ shows a lower cost. Consider the operation related to s_0^* in match M, if $(*, s_0^*) \in M$, it is natural to remove s_0^*, since it introduces extra cost. In addition, if $(t_i, s_0^*) \in M$, since $\Delta_m(t_i, s_0^*) = |t_i - s_0^*| \geq |t_0 - s_0^*| > (\lambda_a + \lambda_d)$, we can just replace it with $(t_i, *)$ and $(*, s_0^*)$, which only generate $(\lambda_a + \lambda_d)$ cost. To conclude, we prove that $t_0 - (\lambda_a + \lambda_d) \leq s_0^* \leq t_{n-1}$. □

A.4 Proof of Lemma 4

Proof. (1) Let insert and delete operations $(*, s_b), (t_a, *) \in M$ be in the same period between $(t_i^m, s_i^m), (t_{i+1}^m, s_{i+1}^m) \in M^m$. Without loss of generality, we assume $s_b < t_a$ (otherwise, we can reverse the sequences). Then we consider two scenarios:

(1.a) If $t_a < s_{i+1}^m$, since $s_b < t_a < s_{i+1}^m$ and $\mathbf{s}$ is the regular interval time series, there must exist $s_b \leq s_x < s_{i+1}^m$ that $|s_x - t_a| < \epsilon_T$ and s_x is a delete point. Therefore, we can safely replace $(*, s_x), (t_a, *)$ with (t_a, s_x) which generates lower cost, since $|s_x - t_a| < \epsilon_T < \lambda_a + \lambda_d$, and it contradicts the minimal cost assumption.

(1.b) If $t_a \geq s_{i+1}^m$, we therefore get $s_{i+1}^m \leq t_a < t_{i+1}^m$, i.e., $|t_a - s_{i+1}^m| < |t_{i+1}^m - s_{i+1}^m|$. That is, we can replace (t_{i+1}^m, s_{i+1}^m) with (t_a, s_{i+1}^m) for lower cost, which also contradicts the assumption.

In summary, we prove that the insert and delete operations could not appear simultaneously in adjacent move operations.

(2) If k insert operations are at the start of the sequence, i.e, $(*, s_0), (*, s_1), \ldots, (*, s_{k-1}) \in M$, we can directly remove these insert operations from M and move the start s_0 to s_k, thus reduce the cost by $k\lambda_a$, which contradicts the minimal cost assumption. □

A.5 Proof of Lemma 5

Proof. First, for any interval ϵ_{a+i} of t_{a+i-1} and t_{a+i}, following formula holds:

$$\begin{aligned}
|\epsilon_{a+i} - \epsilon_T| &= |(t_{a+i} - t_{a+i-1}) - (s_{b+i} - s_{b+i-1})| \\
&= |(t_{a+i} - s_{b+i}) - (t_{a+i-1} - s_{b+i-1})| \\
&\leq |t_{a+i-1} - s_{b+i-1}| + |t_{a+i} - s_{b+i}|
\end{aligned}$$

Therefore, by accumulate $|\epsilon_{a+i} - \epsilon_T|$ in the above formula, we have

$$\begin{aligned}
\sum_{1 \leq i \leq k-1} |\epsilon_{a+i} - \epsilon_T| &\leq \sum_{1 \leq i \leq k-1} |t_{a+i-1} - s_{b+i-1}| + |t_{a+i} - s_{b+i}| \\
&= \sum_{1 \leq i \leq k-1} (|t_{a+i-1} - s_{b+i-1}| + |t_{a+i} - s_{b+i}|)
\end{aligned}$$

i.e., Lemma 5 is proved. That is to say, we successfully obtain a bound of the cost by move operations, related to ϵ_T. Figure 4(a) illustrates an example. □

A.6 Proof of Lemma 6

Proof. Since we have k insert points from s_{b+1} to s_{b+k}, as shown in Figure 4(b), by mapping $s_{b+1}, s_{b+2}, \ldots, s_{b+k}$ to $\mathbf{t}$, we could divide the interval ϵ_{a+1} into $k+1$ parts, according to the sub-sequence $t_a, s_{b+1}, s_{b+2}, \ldots, s_{b+k}, t_{a+1}$. Let ϵ^L and

ϵ^R denote the first and last interval respectively in the aforesaid sub-sequence, i.e., $\epsilon^L = s_{b+1} - t_a$, $\epsilon^R = t_{a+1} - s_{b+k}$, we have $\epsilon_{a+1} = \epsilon^L + (k-1)\epsilon_T + \epsilon^R$. Therefore, we have

$$\begin{aligned}|\epsilon_{a+1} - \epsilon_T| &= |\epsilon^L + (k-2)\epsilon_T + \epsilon^R| \\ &= |(s_{b+1} - t_a) + (k-2)\epsilon_T + (t_{a+1} - s_{b+k})| \\ &= |(\epsilon_T + s_b - t_a) + (k-2)\epsilon_T + (\epsilon_T + t_{a+1} - s_{b+k-1})| \\ &= |s_b - t_a + (k+1)\epsilon_T + t_{a+1} - s_{b+k-1}| \\ &\le |s_b - t_a| + |(k+1)\epsilon_T| + |t_{a+1} - s_{b+k-1}| \\ &\le |t_a - s_b| + 2k\lambda_a + |t_{a+1} - s_{b+k-1}|\end{aligned}$$

thus prove Lemma 6. That is to say, we successfully obtain a bound of the cost by insert operations, related to ϵ_T. □

A.7 Proof of Lemma 7

Proof. To prove the bound, we first prove that $s_b < t_{a+1} < \cdots < t_{a+k} < s_{b+1}$. This is obvious, since if $t_{a+1} \le s_b$, we have $t_a < t_{a+1} \le s_b$, i.e., we could find a better match with lower cost by deleting t_a and match t_{a+1} with s_b instead, which contradicts the condition of the minimal cost. Similarity, we can also prove $t_{a+k} < s_{b+1}$, thus we have $s_b < t_{a+1} < \cdots < t_{a+k} < s_{b+1}$. That is to say, if $k \ge 2$, for $2 \le i \le k$, we have $\epsilon_{a+i} = t_{a+i} - t_{a+i-1} < s_{b+1} - s_b < \epsilon_T$. Next, for $|\epsilon_{a+1} - \epsilon_T|$, we have the following result:

$$\begin{aligned}|\epsilon_{a+1} - \epsilon_T| &= |t_{a+1} - t_a - (s_{b+1} - s_b)| \\ &\le |s_{b+1} - t_{a+1}| + |t_a - s_b| \\ &< |s_{b+1} - s_b| + |t_a - s_b| \\ &= \epsilon_T + |t_a - s_b|\end{aligned}$$

For $|\epsilon_{a+k+1} - \epsilon_T|$, similar result also holds:

$$\begin{aligned}|\epsilon_{a+k+1} - \epsilon_T| &= |t_{a+k+1} - t_{a+k} - (s_{b+1} - s_b)| \\ &\le |t_{a+k+1} - s_{b+1}| + |t_{a+k} - s_b| \\ &< |t_{a+k+1} - s_{b+1}| + |s_{b+1} - s_b| \\ &= \epsilon_T + |t_{a+k+1} - s_{b+1}|\end{aligned}$$

Combining the aforesaid formulas, we finally prove Lemma 7:

$$\begin{aligned}\sum_{1\le i\le k+1} |\epsilon_{a+i} - \epsilon_T| &< |\epsilon_{a+1} - \epsilon_T| + (k-1)\epsilon_T + |\epsilon_{a+k+1} - \epsilon_T| \\ &\le |t_a - s_b| + (k+1)\epsilon_T + |t_{a+k+1} - s_{b+1}| \\ &\le |t_a - s_b| + 2k\epsilon_T + |t_{a+k+1} - s_{b+1}| \\ &\le |t_a - s_b| + 2k\lambda_d + |t_{a+k+1} - s_{b+1}|\end{aligned}$$

An example of delete bound is presented in Figure 4(c). □

A.8 Proof of Proposition 8

Proof. (1) First, assume that $\mathbf{t}$ does not include delete points at the start or the end, and $\mathbf{s}$ is the best target sequence for $\mathbf{t}$ with match M. Analogous to Lemma 4, consider a subsequence of M that include all the move operations in time order, denoted as $M^m = \{(t_0^m, s_0^m), (t_1^m, s_1^m), \ldots (t_{p-1}^m, s_{p-1}^m)\} \subseteq M$ of size p, we can there divide M into $p-1$ parts, denoted as $M_1, M_2, \ldots, M_{p-1}$, where $M_i = \{(x,y) | x \in [t_{i-1}^m, t_i^m] \cup \{*\}, y \in [s_{i-1}^m, s_i^m] \cup \{*\}, (x,y) \in M\}$. For simplicity, let $\Delta(M_i) = \Delta(\mathbf{t}, \mathbf{s}, M_i)$ and $\epsilon_i^m = t_{i+1}^m - t_i^m$, we therefore discuss $\Delta(M_i)$ according to the types of operations:

(a) if $\Delta(M_i)$ does not contain insert or delete operations, $\Delta(M_i) = |s_{i-1}^m - t_{i-1}^m| + |s_i^m - t_i^m| \ge |\epsilon_i^m - \epsilon_T|$ (Lemma 5).

(b) if $\Delta(M_i)$ contains k insert operations, $\Delta(M_i) = |s_{i-1}^m - t_{i-1}^m| + |s_i^m - t_i^m| + k\lambda_a \ge |\epsilon_i^m - \epsilon_T| - k\lambda_a$ (Lemma 6).

(c) if $\Delta(M_i)$ contains k delete operations, $\Delta(M_i) = |s_{i-1}^m - t_{i-1}^m| + |s_i^m - t_i^m| + k\lambda_d \ge \sum_{1\le i\le k+1} |\epsilon_{a+i} - \epsilon_T| - k\lambda_d$, where a denotes the index of t_i^m in $\mathbf{t}$. (Lemma 7).

Accumulate $\Delta(M_i)$ together, we have

$$\begin{aligned}\sum \Delta(M_i) &= \sum_{1\le i\le p} (|s_{i-1}^m - t_{i-1}^m| + |s_i^m - t_i^m|) \\ &\quad + \sum_{(*,s_j)\in M} \lambda_a + \sum_{(t_i,*)\in M} \lambda_b \\ &\le 2 \sum_{(t_i,s_j)\in M} |s_j - t_i| + \sum_{(*,s_j)\in M} \lambda_a + \sum_{(t_i,*)\in M} \lambda_b \\ &= 2\Delta(\mathbf{t}, \mathbf{s}, M) - \sum_{(*,s_j)\in M} \lambda_a - \sum_{(t_i,*)\in M} \lambda_b\end{aligned}$$

In the meantime, by considering (1),(2),(3), we have

$$\sum \Delta(M_i) \ge \sum_{i=1}^{n-1} |\epsilon_T - \epsilon_i| - \sum_{(*,s_j)\in M} \lambda_a - \sum_{(t_i,*)\in M} \lambda_b$$

By combining and reorganizing the formulas, we finally get

$$\Delta(\mathbf{t}, \mathbf{s}, M) \ge \frac{\sum_{i=1}^{n-1} |\epsilon_T - \epsilon_i|}{2} \tag{16}$$

(2) Second, if $\mathbf{t}$ contains k delete points $t_0, t_1, \ldots, t_{k-1}$ at the start, since they result in $k\lambda_d$ cost into the overall cost, for the first k points $t_0, t_1, \ldots, t_{k-1}$ in $\mathbf{t}$, we thus have

$$\frac{\sum_{i=1}^{k-1} \min(|\epsilon_T - \epsilon_i|, \lambda_d)}{2} \le \sum_{i=1}^{k-1} \min(|\epsilon_T - \epsilon_i|, \lambda_d) \le k\lambda_d \tag{17}$$

For the other $n-k$ points $\{t_k, t_{k+1}, \ldots, t_{n-1}\}$ in $\mathbf{t}$, denoted by $\mathbf{t}'$, following Formulas 16, we have $\Delta(\mathbf{t}', \mathbf{s}, M) \ge \frac{\sum_{i=k}^{n-1} |\epsilon_T - \epsilon_i|}{2}$. Combining with Formula 17, we thus have

$$\Delta(\mathbf{t}, \mathbf{s}, M) \ge k\lambda_d + \Delta(\mathbf{t}', \mathbf{s}, M) \ge \frac{\sum_{i=1}^{n-1} \min(|\epsilon_T - \epsilon_i|, \lambda_d)}{2} \tag{18}$$

(3) Finally, if $\mathbf{t}$ contains k delete points at the end, the same formula as Formula 18 could be proved by reversing the time series.

Therefore, considering all the three scenarios, by combining Formulas 16 and 18, we finally prove that

$$\Delta(\mathbf{t}, \mathbf{s}, M) \ge \frac{\sum_{i=1}^{n-1} \min(|\epsilon_T - \epsilon_i|, \lambda_d)}{2} \tag{19}$$

□

A.9 Proof of Corollary 9

Proof. Since $|\epsilon_T - \epsilon_i| \le \lambda_d$, $\frac{\sum_{i=1}^{n-1} \min(|\epsilon_T - \epsilon_i|, \lambda_d)}{2} = \frac{\sum_{i=1}^{n-1} |\epsilon_T - \epsilon_i|}{2}$, we skip the proof here since the monotonicity of the above formula is obvious. □

A.10 Proof of Corollary 10

Proof. Since $\Delta(t_i, *) = \lambda_d, i = 0, 1, \ldots, d$, by dividing $\mathbf{t}$ into 2 parts at t_d and employ Proposition 8 on $\{t_{d+1}, \ldots, t_n\}$, it is easy to show the correctness of Corollary 10. □

A.11 Proof of Corollary 11

Proof. Let $g(d)$ denote the lower bound of $\Delta(\mathbf{t}, \mathbf{s}, M)$, i.e., $g(d) = d\lambda_d + \frac{\sum_{i=d+1}^{n-1} \min(|\epsilon_T - \epsilon_i|, \lambda_d)}{2}, i = 0, 1, \ldots, d$. We have $g(d) - g(d-1) = \lambda_d - \frac{\min(|\epsilon_T - \epsilon_d|, \lambda_d)}{2} \geq \frac{\lambda_d}{2} > 0$. Therefore, $g(d)$ is monotonically increasing. □

A.12 Proof of Proposition 12

Proof. By leveraging Proposition 1 twice on $dp^L(i, j)$ and $dp^R(i, j)$, the correctness of the bi-directional method is proved. □

A.13 Proof of Lemma 13

Proof. Let M_{ab} denote the match that minimizes $\Delta(\mathbf{t}^a, \mathbf{t}^b, M)$, i.e., $\Delta(\mathbf{t}^a, \mathbf{t}^b) = \min_M \Delta(\mathbf{t}^a, \mathbf{t}^b, M) = \Delta(\mathbf{t}^a, \mathbf{t}^b, M_{ab})$. Analogously, let M_{ac}, M_{cb} minimize $\Delta(\mathbf{t}^a, \mathbf{t}^c, M)$ and $\Delta(\mathbf{t}^c, \mathbf{t}^b, M)$, respectively. We will then construct an M'_{ab} based on M_{ac} and M_{cb}, such that $\Delta(\mathbf{t}^a, \mathbf{t}^b, M'_{ab}) \leq \Delta(\mathbf{t}^a, \mathbf{t}^c, M_{ac}) + \Delta(\mathbf{t}^c, \mathbf{t}^b, M_{cb})$.

According to M_{ac} and M_{cb}, for each element in M_{ac} and M_{cb}, we consider all six scenarios that are possible, summarized in the table below. For instance, in Scenario 1, if $(t_i^a, t_j^c) \in M_{ac} \wedge (t_j^c, t_k^b) \in M_{cb}$, i.e., t_i^a is moved to t_j^c in M_{ac} and then moved to t_k^b in M_{cb}, we add (t_i^a, t_k^b) to M'_{ab}. For the other five scenarios, analogously, we add corresponding elements to M'_{ab} (except Scenario 5, where we have no need to update M'_{ab}). According to the comparison of the last two columns in the table, in all scenarios, we have $\Delta(\mathbf{t}^a, \mathbf{t}^c, M_{ac}) + \Delta(\mathbf{t}^c, \mathbf{t}^b, M_{cb}) \geq \Delta(\mathbf{t}^a, \mathbf{t}^b, M'_{ab})$. We can therefore conclude that, $\Delta(\mathbf{t}^a, \mathbf{t}^c, M_{ac}) + \Delta(\mathbf{t}^c, \mathbf{t}^b, M_{cb}) \geq \Delta(\mathbf{t}^a, \mathbf{t}^b, M'_{ab})$.

To this end, we construct M'_{ab} based on M_{ac} and M_{cb}, and obtain $\Delta(\mathbf{t}^a, \mathbf{t}^b, M'_{ab}) \leq \Delta(\mathbf{t}^a, \mathbf{t}^c, M_{ac}) + \Delta(\mathbf{t}^c, \mathbf{t}^b, M_{cb})$. Since M_{ab} minimizes $\Delta(\mathbf{t}^a, \mathbf{t}^b, M)$, i.e., $\Delta(\mathbf{t}^a, \mathbf{t}^b, M_{ab}) \leq \Delta(\mathbf{t}^a, \mathbf{t}^b, M'_{ab})$, we prove that $\Delta(\mathbf{t}^a, \mathbf{t}^b, M_{ab}) \leq \Delta(\mathbf{t}^a, \mathbf{t}^c, M_{ac}) + \Delta(\mathbf{t}^c, \mathbf{t}^b, M_{cb})$, i.e., $\Delta(\mathbf{t}^a, \mathbf{t}^b) \leq \Delta(\mathbf{t}^a, \mathbf{t}^c) + \Delta(\mathbf{t}^c, \mathbf{t}^b)$. □

A.14 Proof of Proposition 14

Proof. Let $s_{md}^{appr}, s_{md}^{exact}$ denote the median of $\mathbf{s}^{appr}$ and $\mathbf{s}^{exact}$, respectively. According to Algorithm 4, we have $s_{md}^{appr} = t_{md}$, and the interval ϵ_T^{appr} of $\mathbf{s}^{appr}$ is the median of all the intervals of $\mathbf{t}$, i.e., ϵ_{md}. Let $\mathbf{s}'$ denote an arbitrary length regular interval time series, having the same interval and median as $\mathbf{s}^{appr}$, i.e., $s'_{md} = t_{md}$ and the interval of $\mathbf{s}'$ is also ϵ_{md}. Since $\mathbf{s}^{appr}$ is computed by dynamic programming with the minimum cost, according to Proposition 12, we have $\Delta(\mathbf{t}, \mathbf{s}^{appr}) \leq \Delta(\mathbf{t}, \mathbf{s}')$. Combining with the triangle inequation in Lemma 13, we have

$$\Delta(\mathbf{t}, \mathbf{s}^{appr}) \leq \Delta(\mathbf{t}, \mathbf{s}') \leq \Delta(\mathbf{t}, \mathbf{s}^{exact}) + \Delta(\mathbf{s}^{exact}, \mathbf{s}')$$

holding for an any $\mathbf{s}'$, i.e., $\Delta(\mathbf{t}, \mathbf{s}^{appr}) - \Delta(\mathbf{t}, \mathbf{s}^{exact}) \leq \Delta(\mathbf{s}^{exact}, \mathbf{s}')$. Recall that $\Delta(\mathbf{t}, \mathbf{s}^{appr})$ and $\Delta(\mathbf{t}, \mathbf{s}^{exact})$ are the costs of approximate and exact methods, respectively. That is, the difference of the approximate algorithm to the optimal solution is bounded by $\Delta(\mathbf{s}^{exact}, \mathbf{s}')$. Next, by dividing $\Delta(\mathbf{t}, \mathbf{s}^{exact})$ on both sides of the aforesaid formula, we derive

$$\frac{\Delta(\mathbf{t}, \mathbf{s}^{appr})}{\Delta(\mathbf{t}, \mathbf{s}^{exact})} \leq 1 + \frac{\Delta(\mathbf{s}^{exact}, \mathbf{s}')}{\Delta(\mathbf{t}, \mathbf{s}^{exact})}$$

We therefore further discuss the range of $\frac{\Delta(\mathbf{s}^{exact}, \mathbf{s}')}{\Delta(\mathbf{t}, \mathbf{s}^{exact})}$, to get the bound for the approximate algorithm.

For the range of $\Delta(\mathbf{s}^{exact}, \mathbf{s}')$, one possible solution is to delete all the points in $\mathbf{s}^{exact}$ and let $|\mathbf{s}'| = 0$. Such a solution has the cost of $m\lambda_d$, where m is the length of $\mathbf{s}^{exact}$, we thus have $\Delta(\mathbf{s}^{exact}, \mathbf{s}') \leq m\lambda_d$. Next, we consider the range of m. Lemma 2 provides the upper bound for m. We consider again the possible solution of deleting all the points in $\mathbf{t}$, which has cost $n\lambda_d$. Therefore, according to Lemma 2, we have $m \leq \lfloor \frac{c}{\lambda_a} \rfloor + n = \lfloor \frac{n\lambda_d}{\lambda_a} \rfloor + n$. Following the setting of $\lambda_a = \lambda_d$ in Section 3.4.6, it derives $m \leq 2n$. Combining $\Delta(\mathbf{s}^{exact}, \mathbf{s}') \leq m\lambda_d$ and $m \leq 2n$, we thus have $\Delta(\mathbf{s}^{exact}, \mathbf{s}') \leq 2n\lambda_d$.

For the range of $\Delta(\mathbf{t}, \mathbf{s}^{exact})$, referring to Proposition 8, we have $\Delta(\mathbf{t}, \mathbf{s}^{exact}) \geq \frac{\sum_{i=1}^{n-1} \min(|\epsilon_T - \epsilon_i|, \lambda_d)}{2}$. Since the minimum value of $\frac{\sum_{i=1}^{n-1} \min(|\epsilon_T - \epsilon_i|, \lambda_d)}{2}$ could be obtained by setting $\epsilon_T = \epsilon_{md}$, we have $\Delta(\mathbf{t}, \mathbf{s}^{exact}) \geq \frac{\sum_{i=1}^{n-1} \min(|\epsilon_{md} - \epsilon_i|, \lambda_d)}{2}$. Let $\delta_\epsilon^{min} = \min(|\epsilon_{md} - \epsilon_i|, \lambda_d)$, $i = 1, 2, \ldots, n-1$ and $\epsilon_i \neq \epsilon_{md}$, i.e., δ_ϵ^{min} is the minimum distance of ϵ_i to ϵ_{md}, less than λ_d. It derives $\frac{\sum_{i=1}^{n-1} \min(|\epsilon_{md} - \epsilon_i|, \lambda_d)}{2} \geq \frac{(n-2)\delta_\epsilon^{min}}{2}$ ($n-2$ is to remove $\epsilon_{md} - \epsilon_i = 0$ when m is odd). If $n \leq 2$, the problem is trivial, since the input time series is already regular. If $n > 2$, we have $(n-2) \geq \frac{n}{2}$ and thus $\frac{(n-2)\delta_\epsilon^{min}}{2} \geq \frac{n\delta_\epsilon^{min}}{4}$. Finally, by combining the ranges of $\Delta(\mathbf{s}^{exact}, \mathbf{s}')$ and $\Delta(\mathbf{t}, \mathbf{s}^{exact})$ above, the error bound for the approximate algorithm is given:

$$\frac{\Delta(\mathbf{t}, \mathbf{s}^{appr})}{\Delta(\mathbf{t}, \mathbf{s}^{exact})} \leq 1 + \frac{\Delta(\mathbf{s}^{exact}, \mathbf{s}')}{\Delta(\mathbf{t}, \mathbf{s}^{exact})} \leq 1 + \frac{2n\lambda_d}{\frac{n\delta_\epsilon^{min}}{4}} = 1 + \frac{8\lambda_d}{\delta_\epsilon^{min}}$$

where $\delta_\epsilon^{min} = \min(|\epsilon_{md} - \epsilon_i|, \lambda_d)$, $i = 1, 2, \ldots, n-1$ and $\epsilon_i \neq \epsilon_{md}$. □

A.15 Proof of Corollary 15

Proof. Recall that in the proof of Proposition 14, we have $\Delta(\mathbf{t}, \mathbf{s}^{appr}) - \Delta(\mathbf{t}, \mathbf{s}^{exact}) \leq \Delta(\mathbf{s}^{exact}, \mathbf{s}')$ holding for an any $\mathbf{s}'$. We thereby discuss another choice of $\mathbf{s}'$. Given $x = |\epsilon_T^{exact} - \epsilon_{md}|$, $y = |s_{md}^{exact} - t_{md}|$, another solution is to make $\mathbf{s}'$ the same size as $\mathbf{s}^{exact}$ and move all the points from $\mathbf{s}'$ to $\mathbf{s}^{exact}$ according to time order. We construct the solution as follows: (1) For odd m, we first make it with regular intervals of ϵ_T^{exact} on both sides of s'_{md}. On each side, it takes the cost of $x + 2x + \cdots + \frac{m-1}{2}x = \frac{m^2-1}{8}x$. Next, we move all points of $\mathbf{s}'$ to $\mathbf{s}^{exact}$, with my cost. Therefore, the total cost is $\frac{m^2-1}{4}x + my$. (2) For even m, analogously, we first make it with regular intervals of ϵ_T^{exact} on each both side of s'_{md} with cost $\frac{m^2}{8}x$. Next, we move $\mathbf{s}'$ to $\mathbf{s}^{exact}$, with my cost. Therefore, the total cost is $\frac{m^2}{4}x + my$. Considering both scenarios together, we thus conclude that $\Delta(\mathbf{t}, \mathbf{s}^{appr}) - \Delta(\mathbf{t}, \mathbf{s}^{exact}) \leq \Delta(\mathbf{s}^{exact}, \mathbf{s}') \leq \frac{m^2}{4}x + my$.

If we have $\epsilon_T^{exact} = \epsilon_{md}$, $s_{md}^{exact} = t_{md}$, then $x = |\epsilon_T^{exact} - \epsilon_{md}| = 0$, $y = |s_{md}^{exact} - t_{md}| = 0$. It follows $\Delta(\mathbf{t}, \mathbf{s}^{appr}) - \Delta(\mathbf{t}, \mathbf{s}^{exact}) \leq \frac{m^2x}{4} + my = 0$, i.e., $\Delta(\mathbf{t}, \mathbf{s}^{appr}) = \Delta(\mathbf{t}, \mathbf{s}^{exact})$. □

A.16 Proof of Proposition 16

Proof. To prove Proposition 16, we need to demonstrate that the state transition equation (Formula 7) accurately accounts for the minimum repair cost of the time series formed by considering all possible sub-sequence splits, with any timestamp

t_i as the end point of the sequence. Here, $C[a,b][0]$ represents the minimum repair cost to convert the sub-sequence $\mathbf{t}_{[a,b]}$ to $\mathbf{s}_{[a,b]}$.

(1) When $i = 0$, the equation simplifies to $R_{seg}[0] = C[0,0][0]$. Since $\mathbf{t}_{[0,0]}$ and $\mathbf{s}_{[0,0]}$ are single timestamp, the minimum repair cost is directly given by $C[0,0][0]$.

(2) Assume that the state transition holds for $R_{seg}[k]$, where $k < i$. We will prove that it also holds for $R_{seg}[i]$. Consider the two terms in the minimum function of the state transition equation. (a) For the term $C[0,i][0]$, it considers the minimum repair cost of transforming the segment $\mathbf{t}_{[0,i]}$ into $\mathbf{s}_{[0,i]}$ as a single segment. If the optimal solution for $R_{seg}[i]$ is a single-segment transformation, the answer will be obtained here. (b) For $\min_{j\in(0,i-1)}(R_{seg}[j]+\mathrm{C}[j+1,i][0]))$, this term divides the sequence into two parts: $\mathbf{t}_{[0,j]}$ and $\mathbf{t}_{[j+1,i]}$. According to the assumption, the cost $R_{seg}[j]$ gives the minimum repair cost for the first part. $C[j+1,i][0]$ provides the minimum cost for independently repairing the sub-sequence $\mathbf{t}_{[j+1,i]}$. By taking the minimum value of all possible j, it ensures considering any possible segmentation of the time series, thereby capturing the optimal multi-segment transformation.

In conclusion, the state transition equation in Formula 7 integrates the cost of repairing $\mathbf{t}_{[0,i]}$ as a single segment and the cost of all possible segmentation methods for repair, ensuring that $R_{seg}[i]$ always captures the minimum cost, regardless of whether the repair involves a single segment or multiple segments. □

A.17 Proof of Proposition 17

Proof. Assume that this upper bound is inaccurate, then we have:

$$m_i^* > \lfloor \frac{t_{n-1} - t_0}{\epsilon_T^i} + 1 \rfloor$$

Given the time series $\mathbf{t}_{[0,n-1]} = \{t_0, t_1, ..., t_{n-1}\}$ and the repaired time series $\mathbf{s}_{[0,m_i^*-1]} = \{s_0, s_1, ..., s_{m_i^*-1}\}$, the interval ϵ_T^i ensures that the repaired series is regular. Hence, we have $s_j = t_0 + j\epsilon_T^i$, where $j = 0, 1, ..., m_i^* - 1$. The total length covered by $\mathbf{s^i}$ should span from t_0 to t_{n-1} or less. Therefore,

$$s_{m_i^*-1} = t_0 + (m_i^* - 1)\epsilon_T^i \leq t_{n-1}$$

Rearranging the inequality, we obtain:

$$m_i^* - 1 \leq \frac{t_{n-1} - t_0}{\epsilon_T^i}$$

Adding 1 to both sides, we get:

$$m_i^* \leq \frac{t_{n-1} - t_0}{\epsilon_T^i} + 1$$

Since m_i^* must be an integer, we take the floor function:

$$m_i^* \leq \lfloor \frac{t_{n-1} - t_0}{\epsilon_T^i} + 1 \rfloor$$

This contradicts our assumption that $m_i^* > \lfloor \frac{t_{n-1}-t_0}{\epsilon_T^i} + 1 \rfloor$. Therefore, our initial assumption is false, and the upper bound of m_i^* as stated in the proposition is indeed accurate. □

A.18 Proof of Proposition 18

Proof. The greedy matching algorithm yields a feasible solution for transforming the original series $\mathbf{t}$ into $\mathbf{s}$, satisfying the requirements of Problem 1 and Problem 2 in repairing regular interval time series. This solution effectively repairs the entire original time series $\mathbf{t}$ and produces the repaired series $\mathbf{s}$, ensuring all timestamps in $\mathbf{t}$ are considered.

In Formula 11, the algorithm starts from the last timestamp t_{n-1} of series $\mathbf{t}$ and iterative selects the smallest-cost sub-sequence according to $S^*_{seg}[i,j]$, saving each selected subsequence. The algorithm proceeds by traversing forwards from the timestamp just before the starting position of each subsequence until it reaches t_0. At each step of traversal, the selected sub-sequence satisfies a local optimally condition defined by Formula 11.

The matching M obtained for series $\mathbf{s}$ encompasses all timestamps in $\mathbf{t}$, thereby confirming that the solution M achieved through the greedy matching algorithm is a feasible approach for repairing the series. □