Motivation

Subset selection finds a small subset of most informative items from a large ground set to be used for summarization and other inference and learning tasks.

Sequential data, including time-series, such as video and speech, and ordered data, such as text, form a significant part of modern datasets.

- There exist structural dependencies among sequential data, imposed by the underlying dynamic model, that must play a vital role in summarization.
- Existing subset selection methods ignore dynamics, treating data as a bag of randomly permutable items.

Sequential Facility Location

- **Approach:** Introduce transition model among source set items \( p(x_i|x_0, \ldots, x_n) \)
  - Target set has a sequential structure \( Y = \{y_1, \ldots, y_M\} \)
  - \( x_i \) denotes the representative of \( y_i \), for \( i \in \{1, \ldots, T\} \).

  \[
  \max_{\{z_{t,1}, \ldots, z_{t,T}\} \subseteq \{1, \ldots, M\}^T} \Psi(t_1, \ldots, t_T) = \Phi_{\text{enc}}(t_1, \ldots, t_T) + \Phi_{\text{card}}(t_1, \ldots, t_T) + \Phi_{\text{dyn}}(t_1, \ldots, t_T)
  \]

  Encoding: \( \Phi_{\text{enc}}(t_1, \ldots, t_T) = \prod_{t=1}^{T} \phi_{\text{enc}}(t) \prod_{t=1}^{T} \phi_{\text{enc}}(t_1, \ldots, t_T) \)

  Cardinality: \( \Phi_{\text{card}}(t_1, \ldots, t_T) = \exp(-\lambda \|z_{t,1}, \ldots, z_{t,T}\|) \)

  Dynamics: \( \Phi_{\text{dyn}}(t_1, \ldots, t_T) = \prod_{t=1}^{T} \phi_{\text{dyn}}(t_1, \ldots, t_T) \prod_{t=1}^{T} \phi_{\text{dyn}}(t_1, \ldots, t_T) \)

- **Integer Binary Optimization Formulation:**
  - Binary assignment variable \( z_{t,i} \in \{0,1\} \) indicates if \( x_i \) is a representative of \( y_t \).
  - Consider first-order Markov model and maximize log \( \Psi \)

- **Optimization via Max-Sum Message Passing:** cast the optimization as a MAP inference on binary random variables.

Prior Work

Determinantal Process (DPPs) choose subset(s) \( Y \) of data \( X \)

- Markov DPP [1]: successively selects items, diverse from previously selected items.
- SeqDPP [2]: divides a sequence into windows \( Y_t \) and selects sets \( Y_t \) diverse within window and with respect to items selected in previous window.

- **Limitations:** i) do not consider dynamics of data; ii) single set summarization.

Facility Location Review

- Given: source set \( X \), target set \( Y \), and pairwise dissimilarity \( d_{ij} \).
- \( d_{ij} \): how well \( x_i \) represents \( y_j \), smaller means better.
- Goal: find a small subset \( S \subseteq X \) to represent every item \( y_j \in Y \).
- Minimize cardinality plus encoding quality of the representative set:

  \[
  \min_{S \subseteq X} \lambda |S| + \sum_{i,j \in S} d_{ij}
  \]

Contribution

- Develop a sequential subset selection framework incorporating dynamics.
- Form potentials to optimize encoding, cardinality and coherency of the summary.
- Propose a binary optimization over data assignments to representatives.
- Develop a max-sum message passing and an ADMM framework.

Experiments

- **Synthetic Experiments:** \( X \): means of 50 Gaussians, \( Y \): from a Markov model

- **Instructional Video Summarization**
  - Use the Instructional Video dataset [3]: 5 tasks, 30 videos per task available.
  - Fit HMM to training data to construct transition model.
  - Use SeqFL to choose representative HMM states for each test video.
  - Assign labels to states based on training set nearest neighbors.
  - Align sequences of representatives from all test videos to form final summary.

- **Facility Location Review**
  - Change
  - Make coffee
  - CPR
  - Jump car
  - Repeat plant
  - All tasks

- **Facility Location**
  - Task:
  - Ground Truth:
  - SeqFL

- **Facility Location**
  - Task:
  - Ground Truth:
  - SeqFL

- **Facility Location**
  - Task:
  - Ground Truth:
  - SeqFL

- **Facility Location**
  - Task:
  - Ground Truth:
  - SeqFL

- **Facility Location**
  - Task:
  - Ground Truth:
  - SeqFL

Figure 1: Dynamic cost, total cost and diversity score as a function of the number of representatives.

Figure 2: Number of representatives, encoding cost and dynamic cost of SeqFL as a function of the parameters \((\lambda, \beta)\).

Figure 3: Summaries obtained for the task of repairing a plant (top) and performing CPR (bottom).