Target Tracking: Lecture 5

Multiple Target Tracking: Part II

Gustaf Hendeby
hendeby@isy.liu.se

Div. Automatic Control
Dept. Electrical Engineering
Linköping University

December 10, 2014

Lecture Outline

Last Lecture

- Intro multi-target tracking (MTT)
- Single hypothesis tracker (SHT)
  - Global nearest neighbor (GNN)
  - Joint Probabilistic Data Association (JPDA)
- Auction algorithm
- Fundamental theorem of target tracking

Multiple Hypothesis Tracking (MHT)

- MHT: consider multiple associations hypotheses over time
- Started with the conceptual MHT
- Integrated track initialization
- Two principal implementations
  - hypotheses based
  - track based

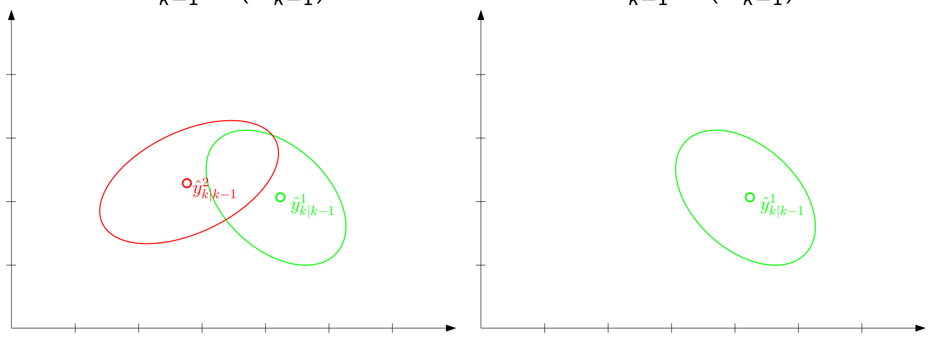
Conceptual MHT: basic idea

- Described in Reid (1979)
- Intuitive hypothesis based *brute force* implementation
- Between consecutive time instants, different association hypotheses, $\{\Theta^i_{k-1}\}_{i=1}^{N_h}$, are kept in memory
- **Idea:** generate all possible hypotheses, and then prune to avoid combinatorial hypotheses growth
- Hypothesis limiting techniques:
  - clustering
  - pruning low probability hypotheses
  - N-scan pruning
  - combining similar hypotheses
  - ...

Representing Hypotheses

Each hypothesis, $\{\Theta^i_{k-1}\}_{i=1}^{N_h}$, is characterized by the *number of targets* (tracks) and their corresponding *sufficient statistics*

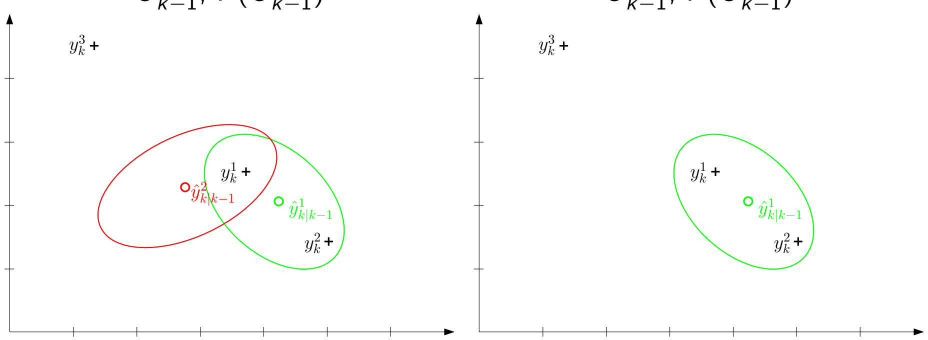
Θ^1_{k-1}, $P(\Theta^1_{k-1})$ Θ^2_{k-1}, $P(\Theta^2_{k-1})$



Representing Hypotheses

Each hypothesis, $\{\Theta^i_{k-1}\}_{i=1}^{N_h}$, is characterized by the *number of targets* (tracks) and their corresponding *sufficient statistics*

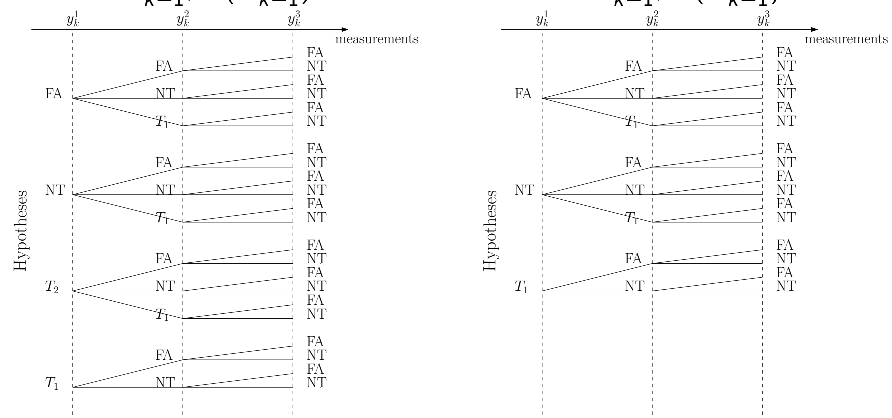
Θ^1_{k-1}, $P(\Theta^1_{k-1})$ Θ^2_{k-1}, $P(\Theta^2_{k-1})$



Generating Hypotheses

Form $\Theta^\ell_k \triangleq \{\theta_k, \Theta^i_{k-1}\}$

Θ^1_{k-1}, $P(\Theta^1_{k-1})$ Θ^2_{k-1}, $P(\Theta^2_{k-1})$



Computing Hypothesis Probabilities

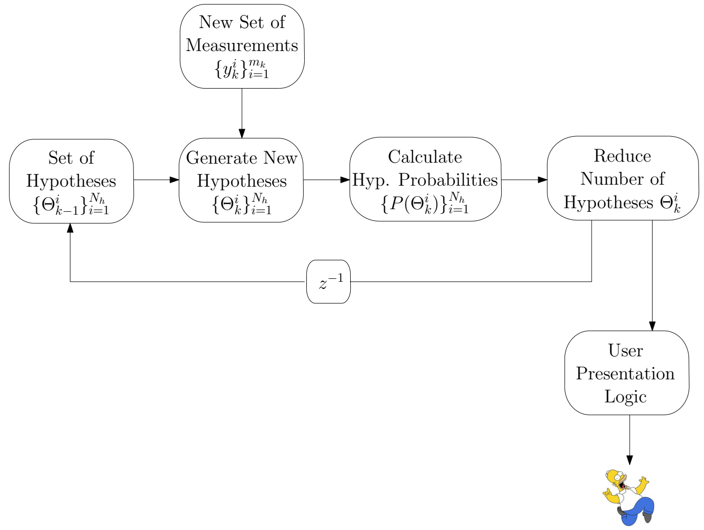
Let $\Theta_k^\ell \triangleq \{\theta_k, \Theta_{k-1}^i\}$, then
(using the "Fundamental Theorem of TT")

$$P(\Theta_k^\ell | y_{0:k}) \propto p(y_k|\Theta_k^\ell, y_{0:k-1}) P(\theta_k|\Theta_{k-1}^i, y_{0:k-1}) P(\Theta_{k-1}^i|y_{0:k-1})$$
$$\propto \beta_{\text{FA}}^{m_k^{\text{FA}}} \beta_{\text{NT}}^{m_k^{\text{NT}}} \Bigg[\prod_{j\in\mathcal{J}_D^i} P_D^j p_{k|k-1}^j(y_k^{\theta_k^{-1}(j)}) \Bigg] \Bigg[\prod_{j\in\mathcal{J}_{ND}^i} (1 - P_D^j P_G^j) \Bigg] P(\Theta_{k-1}^i|y_{0:k-1})$$

Note

The sets $\mathcal{J}_D^i$ and $\mathcal{J}_{ND}^i$ depend on Θ_{k-1}^i! The number of targets and target estimates usually differ between hypotheses.

System Overview

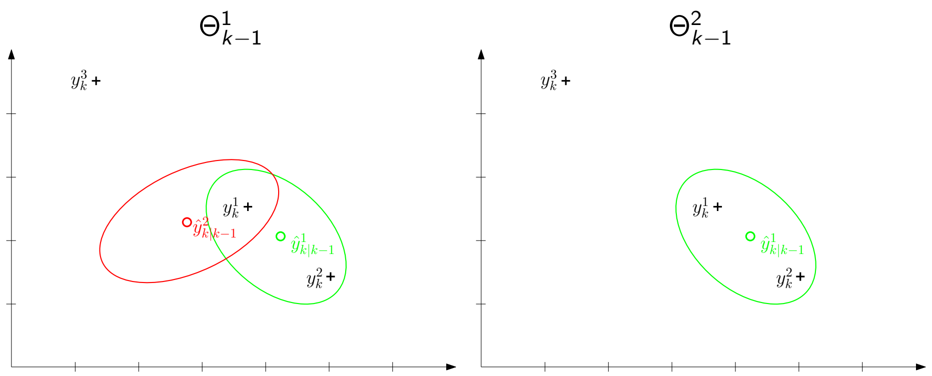


Reducing Complexity

- Clustering
- Pruning of low probability hypotheses
- N-scan pruning
- Merging similar hypotheses

Clustering

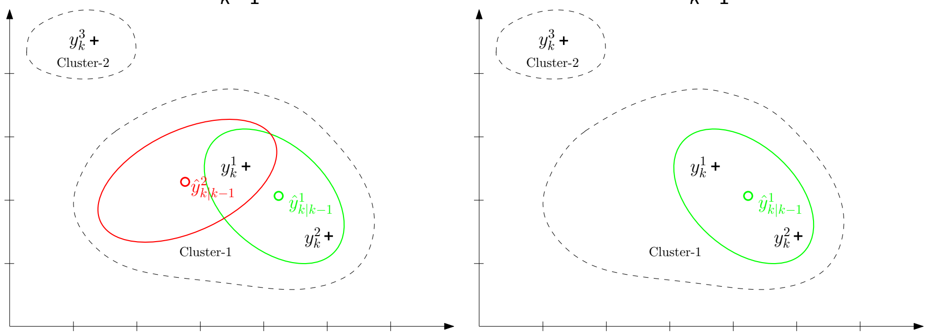
- Group targets without common measurements, and handle the groups separately



Clustering

- Group targets without common measurements, and handle the groups separately

Θ^1_{k-1} Θ^2_{k-1}



Clustering: cluster management

- When targets get closer
  - If measurement falls inside the gates of tracks in different clusters, merge the clusters
  - The hypotheses for each cluster are combined into a super-hypotheses
- When targets separate
  - If a group of tracks cluster do not share measurements with the other tracks in the cluster (for a period of time), split the cluster
  - Hypotheses for the cluster are also divided into smaller hypotheses corresponding to two smaller clusters.

Clustering: process clusters separately (1/2)

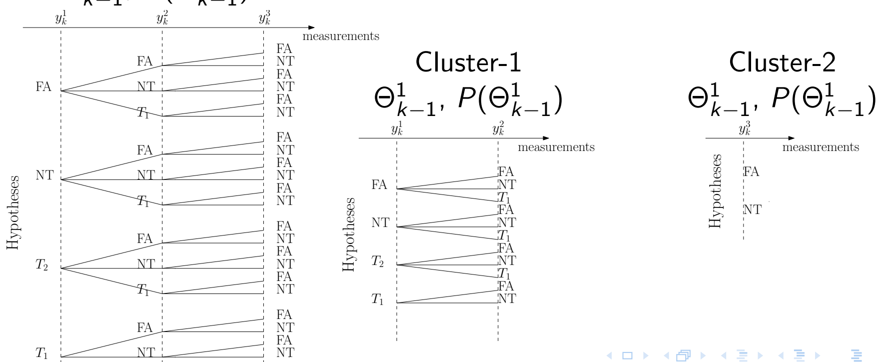
Hypotheses generation

- Form $\Theta^\ell_k \triangleq \{\theta_k, \Theta^i_{k-1}\}$ for each cluster as if the other clusters do not exist.

Without clustering
Θ^1_{k-1}, $P(\Theta^1_{k-1})$

Cluster-1
Θ^1_{k-1}, $P(\Theta^1_{k-1})$

Cluster-2
Θ^1_{k-1}, $P(\Theta^1_{k-1})$



Clustering: process clusters separately (1/2)

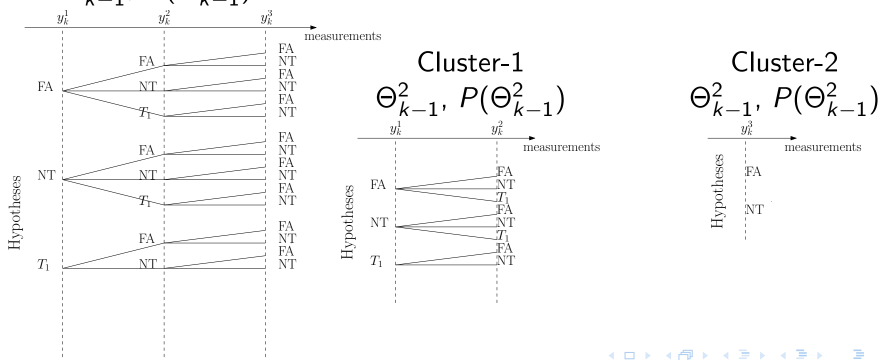
Hypotheses generation

- Form $\Theta^\ell_k \triangleq \{\theta_k, \Theta^i_{k-1}\}$ for each cluster as if the other clusters do not exist.

Without clustering
Θ^2_{k-1}, $P(\Theta^2_{k-1})$

Cluster-1
Θ^2_{k-1}, $P(\Theta^2_{k-1})$

Cluster-2
Θ^2_{k-1}, $P(\Theta^2_{k-1})$



Clustering: process clusters separately (2/2)

Hypotheses reduction

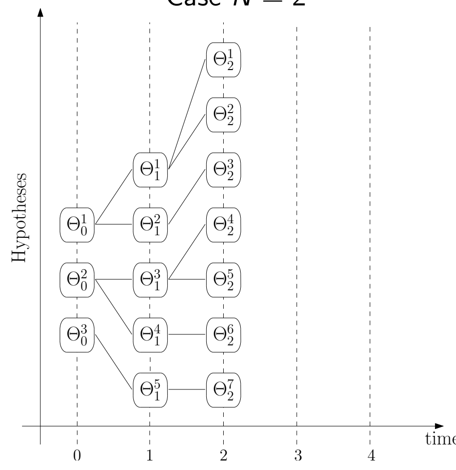
For each cluster:

- Delete hypotheses with probability below a threshold, γ_p (*e.g.*, $\gamma_p = 0.001$)
  $$\text{Deletion Condition: } P(\Theta_k^i) < \gamma_p$$
- Keep only the most probable hypotheses with a total probability mass above a threshold, γ_c (*e.g.*, $\gamma_c = 0.99$)
  $$\text{Deletion Condition: } \sum_{k=1}^{i} P(\Theta_k^{\ell_k}) > \gamma_c$$
  where ℓ_k is a sequence such that $P(\Theta_k^{\ell_k}) \geq P(\Theta_k^{\ell_{k+1}})$

N-scan Pruning

- This scheme assumes that any uncertainty is perfectly resolved after N time steps
- It is a general commonsense to choose $N \geq 5$ (situation dependent)
- The N last ancestors of each hypothesis must be stored

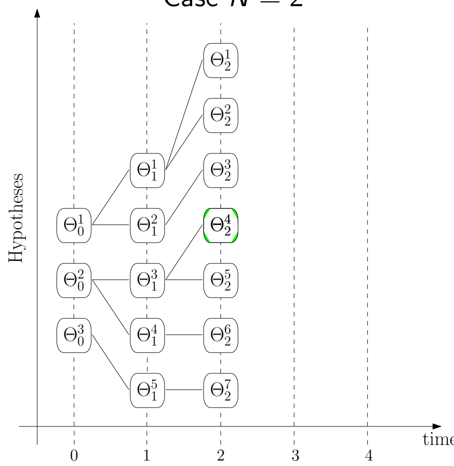
Case $N = 2$



N-scan Pruning

- This scheme assumes that any uncertainty is perfectly resolved after N time steps
- It is a general commonsense to choose $N \geq 5$ (situation dependent)
- The N last ancestors of each hypothesis must be stored

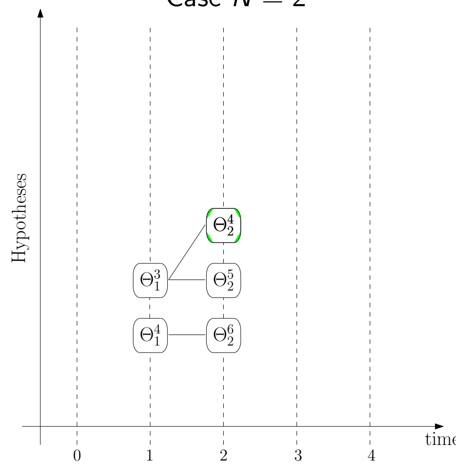
Case $N = 2$



N-scan Pruning

- This scheme assumes that any uncertainty is perfectly resolved after N time steps
- It is a general commonsense to choose $N \geq 5$ (situation dependent)
- The N last ancestors of each hypothesis must be stored

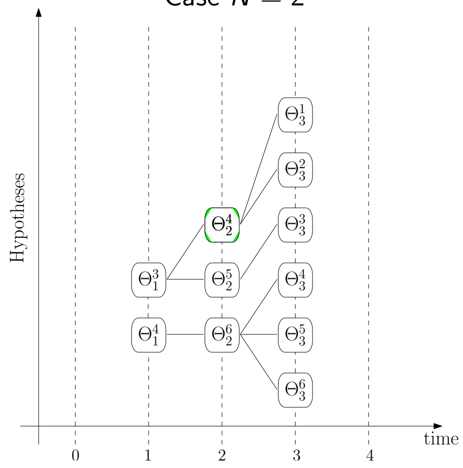
Case $N = 2$



N-scan Pruning

- This scheme assumes that any uncertainty is perfectly resolved after N time steps
- It is a general commonsense to choose $N \geq 5$ (situation dependent)
- The N last ancestors of each hypothesis must be stored

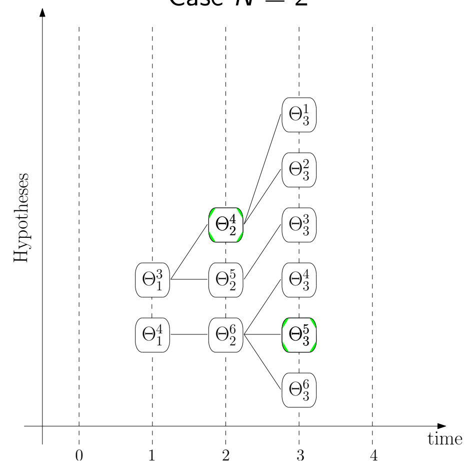
Case $N = 2$



N-scan Pruning

- This scheme assumes that any uncertainty is perfectly resolved after N time steps
- It is a general commonsense to choose $N \geq 5$ (situation dependent)
- The N last ancestors of each hypothesis must be stored

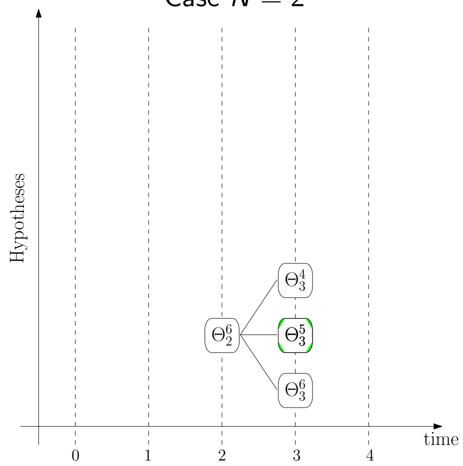
Case $N = 2$



N-scan Pruning

- This scheme assumes that any uncertainty is perfectly resolved after N time steps
- It is a general commonsense to choose $N \geq 5$ (situation dependent)
- The N last ancestors of each hypothesis must be stored

Case $N = 2$



Hypothesis Merging

Reid's original paper suggests to check for hypothesis pairs with:

- the same number of targets (tracks)
- similar track estimates

If these conditions are satisfied:

- merge the hypotheses
- assign the new hypothesis the sum of the combined hypotheses' probability

Summary

- Attractive method since each hypothesis is
  - an alternative representation of reality
  - easily interpreted
- Drawback: generating all possible hypotheses only to discarding (most of) them is inefficient
- Some hypotheses contain the same track; hence fewer unique tracks than hypotheses
- Track based methods were popular until an efficient way to implement a hypothesis based MHT was given by Cox and Hingorani (1996)

Hypothesis-Based MHT

- Proposed by Cox and Hingorani (1996)
- Generate only the best hypotheses, skip hypotheses that will be deleted
- Use the N-best solutions to the assignment problem (introduced last lecture with GNN)
  - *Murty's method*, 1968
- Find the N_h-best hypothesis, generating as few unnecessary hypothesis as possible
- Hypothesis reduction techniques still apply

Assignment Problem: repetition (1/2)

Let $\Theta_k^\ell \triangleq \{\theta_k, \Theta_{k-1}^i\}$.

$$P(\Theta_k^\ell | y_{0:k}) \propto p(y_k | \Theta_k^\ell, y_{0:k-1}) P(\theta_k | \Theta_{k-1}^i, y_{0:k-1}) P(\Theta_{k-1}^i | y_{0:k-1})$$
$$\propto \beta_{\mathrm{FA}}^{m_k^{\mathrm{FA}}} \beta_{\mathrm{NT}}^{m_k^{\mathrm{NT}}} \left[\prod_{j \in \mathcal{J}_D^i} P_D^j p_{k|k-1}^j (y_k^{\theta_k^{-1}(j)}) \right] \left[\prod_{j \in \mathcal{J}_{ND}^i} (1 - P_D^j P_G^j) \right] P(\Theta_{k-1}^i | y_{0:k-1})$$

Divide and multiply the right hand side by

$$C_i \triangleq \prod_{j=1}^{n_T^i} (1 - P_D^j P_G^j) = \prod_{j \in \mathcal{J}_D^i} (1 - P_D^j P_G^j) \prod_{j \in \mathcal{J}_{ND}^i} (1 - P_D^j P_G^j)$$

Assignment Problem: repetition (2/2)

$$P(\Theta_k^\ell | y_{0:k}) \propto \beta_{\mathrm{FA}}^{m_k^{\mathrm{FA}}} \beta_{\mathrm{NT}}^{m_k^{\mathrm{NT}}} \left[\prod_{j \in \mathcal{J}_D^i} \frac{P_D^j p_{k|k-1}^j (y_k^{\theta_k^{-1}(j)})}{1 - P_D^j P_G^j} \right] C_i P(\Theta_{k-1}^i | y_{0:k-1})$$

- $\times$ represents $-\infty$.
- $\ell_{ij} \triangleq \log \frac{P_D^j p_{k|k-1}^j (y_k^i)}{(1 - P_D^j P_G^j)}$.

Logarithmize and form the assignment matrices

$\mathcal{A}_1$	T_1	T_2	FA_1	FA_2	FA_3	NT_1	NT_2	NT_3
y_k^1	ℓ_{11}	ℓ_{12}	$\log \beta_{\mathrm{FA}}$	$\times$	$\times$	$\log \beta_{\mathrm{NT}}$	$\times$	$\times$
y_k^2	ℓ_{21}	$\times$	$\times$	$\log \beta_{\mathrm{FA}}$	$\times$	$\times$	$\log \beta_{\mathrm{NT}}$	$\times$
y_k^3	$\times$	$\times$	$\times$	$\times$	$\log \beta_{\mathrm{FA}}$	$\times$	$\times$	$\log \beta_{\mathrm{NT}}$

$\mathcal{A}_2$	T_1	FA_1	FA_2	FA_3	NT_1	NT_2	NT_3
y_k^1	ℓ_{11}	$\log \beta_{\mathrm{FA}}$	$\times$	$\times$	$\log \beta_{\mathrm{NT}}$	$\times$	$\times$
y_k^2	ℓ_{21}	$\times$	$\log \beta_{\mathrm{FA}}$	$\times$	$\times$	$\log \beta_{\mathrm{NT}}$	$\times$
y_k^3	$\times$	$\times$	$\times$	$\log \beta_{\mathrm{FA}}$	$\times$	$\times$	$\log \beta_{\mathrm{NT}}$

Assignment Problem: N-best solutions

- Given an assignment matrix $\mathcal{A}_i$, the Auction algorithm (or similar) finds the best assignment in polynomial time
- Generalizations of this problem to find the N-best solutions:
  - Formulate as several best assignment problems
  - Solve independently using the Auction algorithm
  - Murty's method

Assignment Problem: Murty's Method

Murty's Method

Given the assignment matrix $\mathcal{A}_i$,

- Find the best solution using Auction algorithm.
- 2nd best solution:
  - Express the 2nd best solution as the solution of a number of best solution assignment problems.
  - Find the solution to each of these problems by Auction.
  - The solution giving the maximum reward (minimum cost) is the second best solution.
- Repeat the procedure for more solutions

Algorithm Outline

- **Aim:** Given hypotheses $\{\Theta^i_{k-1}\}_{i=1}^{N_h}$ and measurements $\{y^i_k\}_{i=1}^{m_k}$, find the N_h best hypotheses $\{\Theta^i_k\}_{i=1}^{N_h}$ (avoid generating all hypotheses)
- **Reminder of Hypothesis Probability**

$$P(\Theta^\ell_k | y_{0:k}) \propto \underbrace{\beta_{\mathrm{FA}}^{m_k^{\mathrm{FA}}} \beta_{\mathrm{NT}}^{m_k^{\mathrm{NT}}} \left[\prod_{j \in \mathcal{J}_D^i} \frac{P_D^j p^j_{k|k-1}(y_k^{\theta_k^{-1}(j)})}{1 - P_D^j P_G^j} \right]}_{\text{Assignment dependent}} \underbrace{C_i P(\Theta^i_{k-1} | y_{0:k-1})}_{\text{Legacy}}$$

- Find $\{\Theta^\ell_k\}_{\ell=1}^{N_h}$ that maximizes $P(\Theta^\ell_k | y_{0:k})$.
- Two steps:
  - Obtain the solution from the assignment (Murty's method)
  - Multiply the obtained quantity by previous hypothesis dependent terms

Generating the N_h-best Hypotheses

Input $\{\Theta^i_{k-1}\}_{i=1}^{N_h}$, $\{P(\Theta^i_{k-1} | y_{0:k-1})\}_{i=1}^{N_h}$, and $\{y^i_k\}_{i=1}^{m_k}$

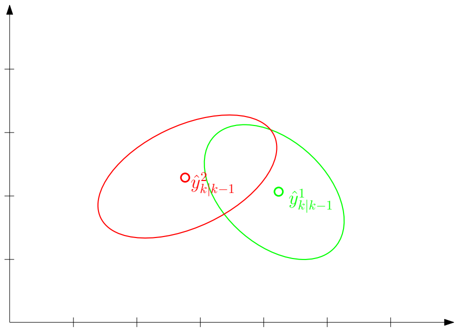
Output HYP-LIST (N hypotheses, decreasing probability)
PROB-LIST (matching probabilities)

1. Initialize all elements in HYP-LIST and PROB-LIST to $\emptyset$ and -1
2. Find assignment matrices $\{\mathcal{A}_i\}_{i=1}^{N_h}$ for $\{\Theta^i_{k-1}\}_{i=1}^{N_h}$
3. For $j = 1 \dots N_h$
   1. For $i = 1 \dots N_h$
      1. For the assignment matrix $\mathcal{A}_i$ find the jth best solution Θ^{ji}_k
      2. Compute the probability $P(\Theta^{ji}_k)$
      3. Update HYP-LIST and PROB-LIST: If the new hypothesis enters the list, discard the least probable entry
      4. If $P(\Theta^{ji}_k)$ is lower than the lowest probability in PROB-LIST discard Θ^{ji}_k and never use $\mathcal{A}_i$ again in subsequent recursions

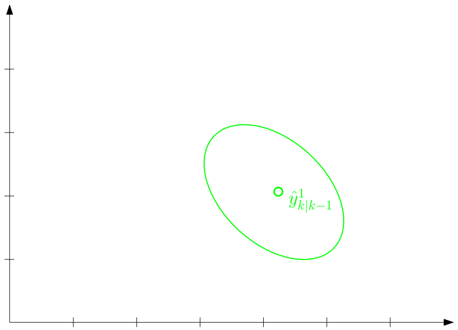
Track-Based MHT: motivation

- Hypotheses usually contain identical tracks — significantly fewer tracks than hypotheses
- **Idea:** Store tracks, T^i, not hypotheses, Θ^i, over time

Θ^1_{k-1}, $P(\Theta^1_{k-1})$

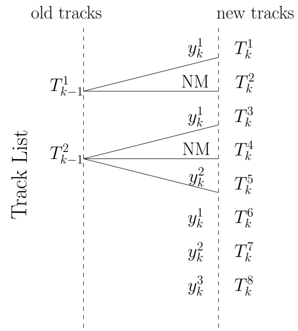


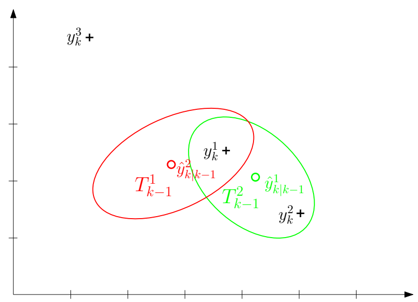
Θ^2_{k-1}, $P(\Theta^2_{k-1})$



Track-Based MHT: principle

- Tracks at time k, $\{T^i_k\}_{i=1}^{N_t}$
- Track scores, $\mathrm{Sc}(T^i_k)$
- Form a track tree, not a hypothesis tree
- Delete tracks with low scores

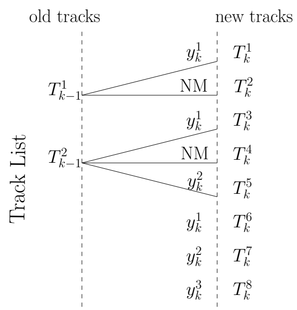




Hypotheses Generation

- Hypothesis: a collection of compatible tracks: $\Theta^1_k = \{T^1_k, T^5_k, T^8_k\}$, $\Theta^2_k = \{T^2_k, T^3_k, T^7_k, T^8_k\}$
- Generating hypothesis is needed for reducing the number of tracks further and for user presentation
- Use only tracks with high score
- Keep track compatibility information (*e.g.*, in a binary matrix)

	T^1_k	T^2_k	T^3_k	T^4_k	T^5_k	T^6_k	T^7_k	T^8_k
T^1_k	0	0	0	1	1	0	1	1
T^2_k		0	1	1	1	1	1	1
T^3_k			0	0	0	0	1	1
T^4_k				0	0	1	1	1
T^5_k					0	1	0	1
T^6_k						0	1	1
T^7_k							0	1
T^8_k								0



Track Scores and Hypotheses Probabilities

- Track probability:

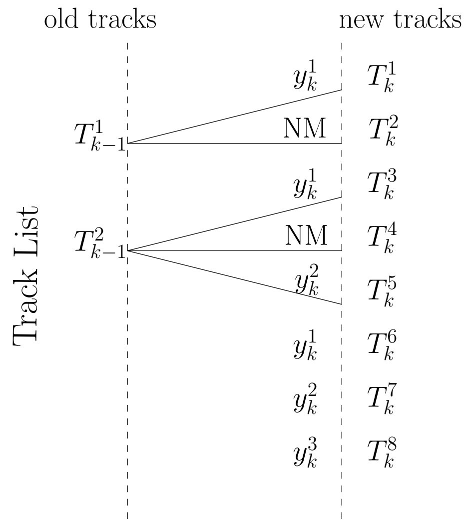
$$P(T^i_k) = \sum_{T^i_k \in \Theta^j_k} P(\Theta^j_k)$$

- Hypothesis score:

$$\mathrm{Sc}(\Theta^i_k) = \sum_{T^j_k \in \Theta^i_k} \mathrm{Sc}(T^j_k)$$

- Hypothesis probability:

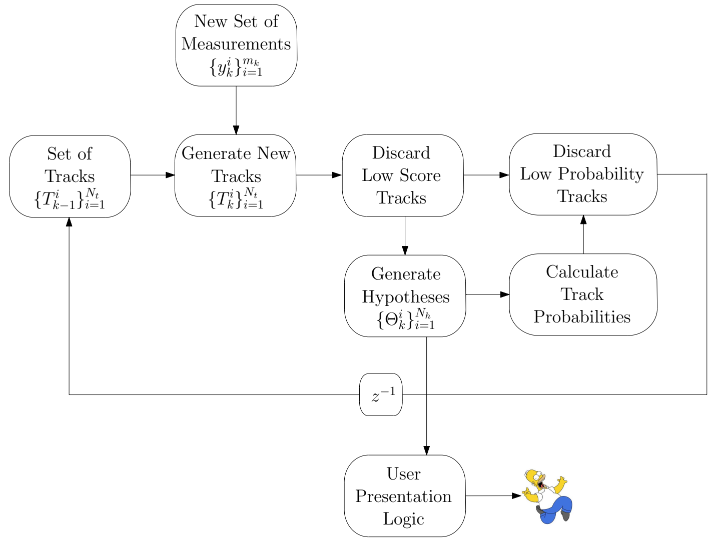
$$P(\Theta^i_k) = \frac{\exp\big(\mathrm{Sc}(\Theta^i_k)\big)}{1 + \sum_{j=1}^{N_h} \exp\big(\mathrm{Sc}(\Theta^j_k)\big)}$$



Complexity Reducing Techniques

- Cluster incompatible tracks for efficient hypothesis generation
- Apply N-scan pruning to the track trees
- Merge tracks with common recent measurement history

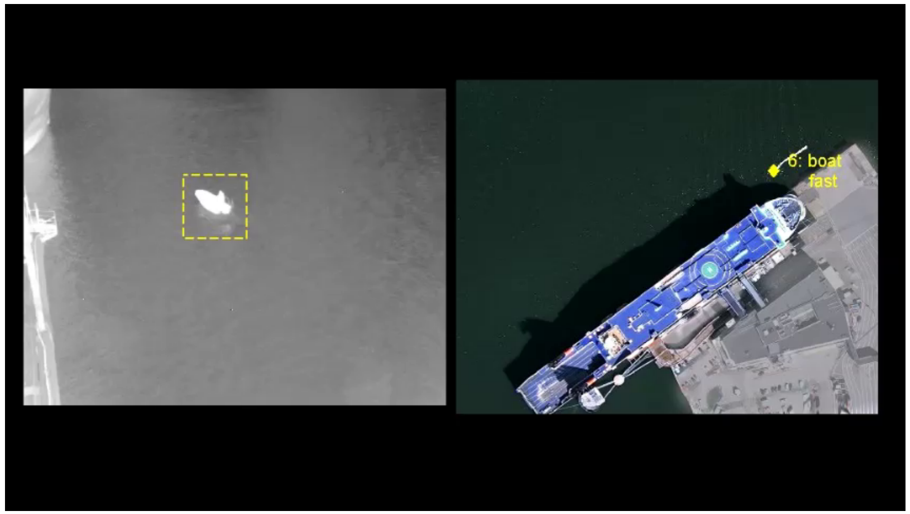
System Components



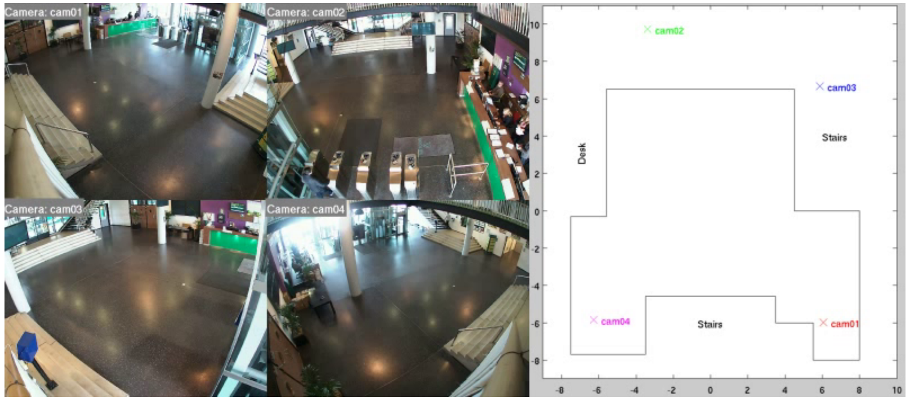
User Presentation Logic

- Maximum probability hypothesis: simplest alternative
  - Possibly jumpy; the maximum probability hypothesis can change erratically
- Show track clusters: (weighted) mean, covariance and expected number of targets
- Keep a separate track list: update at each step with a selection of tracks from different hypotheses
- Consult (Blackman and Popoli, 1999) for details

Example: harbor protection (SUPPORT)



Example: busy indoor environments



Which Multi-TT Method to Use?

Computation \ SNR	Low	Medium	High
Low	Group TT / PHD	GNN	GNN
Medium	MHT	GNN or JPDA	GNN
High	TrBD / MHT	MHT	Any

- GNN and JPDA are very bad in low SNR.
- When using GNN, one generally has to enlarge the overconfident covariances to account for neglected data association uncertainty.
- JPDA has track coalescence and should not be used with closely spaced targets, see the "coalescence avoiding" versions.
- MHT requires significantly higher computational load but it is said to be able to work reasonably under 10–100 times worse SNR.

Learning More (1/2)

Samuel S. Blackman.
Multiple hypothesis tracking for multiple target tracking.
IEEE Transactions on Aerospace and Electronic Systems, 19(1):5–18, January 2004.

Samuel S. Blackman and Robert Popoli.
Design and analysis of modern tracking systems.
Artech House radar library. Artech House, Inc, 1999.
ISBN 1-5853-006-0.

Ingemar J. Cox and Sunita L. Hingorani.
An efficient implementation of Reid's multiple hypothesis tracking algorithm and its evaluation for the purpose of visual tracking.
IEEE Transactions on Pattern Analysis and Machine Intelligence, 18(2):138–150, February 1996.

Ingemar J. Cox and Matthew L. Miller.
On finding ranked assignments with application to multitarget tracking and motion correspondence.
IEEE Transactions on Aerospace and Electronic Systems, 31(1):486–489, January 1995.

Learning More (2/2)

Ingemar J. Cox, Matthew L. Miller, Roy Danchick, and G. E. Newnam.
A comparison of two algorithms for determining ranked assignments with application to multitarget tracking and motion correspondence.
IEEE Transactions on Aerospace and Electronic Systems, 33(1):295–301, January 1997.

Roy Danchick and G. E. Newnam.
Reformulating Reid's MHT method with generalised Murty K-best ranked linear assignment algorithm.
IEE Proceedings-F Radar and Sonar Navigation, 153(1):13–22, February 2006.

Matthew L. Miller, Harold S. Stone, and Ingemar J. Cox.
Optimizing Murty's ranked assignment method.
IEEE Transactions on Aerospace and Electronic Systems, 33(3):851–862, July 1997.

Donald B. Reid.
An algorithm for tracking multiple tragets.
IEEE Transactions on Automatic Control, 24(6):843–854, December 1979.