Future State Prediction for Sensemaking

Rohit Khilnani¹, Soumen Lahiri², Osman Ozaltin³, David L. Roberts¹
¹Computer Science, ²Statistics, ³ISyE
rkkhilna@ncsu.edu, snlahiri@ncsu.edu, oyozalti@ncsu.edu, robertsd@csc.ncsu.edu

Motivation

• Big picture goals:
  • Summarizing time ordered massive data
  • Automatic Identification of actors/key events
  • Prediction of Future State of actors
  • Characterize how the process of summarizing data can affect it’s usefulness for narrative prediction

• Importance: Narrative spaces generated from data can provide intelligence workers with meaningful, easily-interpretable, and actionable information

• We enable more efficient computational inquiry of narrative-focused large data.

Challenges

• Huge data to learn from
  • More than 500 games
  • Average length of game is 40 minutes
  • Parsed game statistics at 30 frames per second
  • About 377,280,000 data snapshots, each having 48 observable attributes
  • Large volumes of noise and irrelevant information contained in data sets

• Analyzing multiple attribute sequences increase problem size even further

• Develop adaptive sampling algorithms and compression methods applicable to different data types including categorical and numerical data

• Develop reliable numerical measures to quantify the loss of information from compression, sampling, and prediction using big data

Approaches and Algorithms

Autoregressive Moving Average Model (ARMA) [1]

• Models time series data with an autoregressive (AR) and moving-average (MA) terms
• Used for analysis and forecasting of future values and can be applied to forecasting character attributes

ARMA Model Predictive Power

• Every 50 time steps, we build a model using data at $t_0$ to $t_i$, test prediction error at time steps $t_{i+1}$ to $t_{i+20}$ and calculate sum of squared error
• This shows us how accurate our model is over time
• As a baseline, we can assume a model that only predicts the mean value of the training set every time
• ARMA model alone doesn’t forecast large changes well

Future Directions

• Modeling both normal gameplay flow and important events
  • Capturing large, frequent changes in character attributes in a forecasting model
  • Predicting major events (e.g. character deaths, large battles, etc.)
  • Incorporating multivariate models using other character attributes
  • Modeling of multiplayer interactions
  • Clustering similar player’s behavior to understand types of players
  • Analyze interaction between groups of players between teams and within teams
• Quantifying uncertainty of both attribute forecasts and event prediction

Example

An example sequence of attribute ‘change in player’s health’:

The sequence would represent an attack being made on the player as health drops after a period of no change

ARMA Model vs. Actual

Error over time for a single hero

Model fails to predict large, frequent changes in health

Results

Our Data – DOTA 2

• Online multiplayer game – Defense of the Ancients 2
  • Two teams of up to five players battle against each other
    • Collaborative, requires coordination
    • Competitive, require combat
  • Each player controls one character in the game
  • Game involves making decisions, collecting items, attacking opponents, helping allies, and has a spatial component
  • Game ends with one team winning over the other by acquiring their home base
  • Provides a reasonable approximation of real-world data streams

Motivation

• Online multiplayer game – Defense of the Ancients 2
  • Two teams of up to five players battle against each other
    • Collaborative, requires coordination
    • Competitive, require combat
  • Each player controls one character in the game
  • Game involves making decisions, collecting items, attacking opponents, helping allies, and has a spatial component
  • Game ends with one team winning over the other by acquiring their home base
  • Provides a reasonable approximation of real-world data streams

Approaches and Algorithms

Autoregressive Moving Average Model (ARMA) [1]

• Models time series data with an autoregressive (AR) and moving-average (MA) terms
• Used for analysis and forecasting of future values and can be applied to forecasting character attributes

ARMA Model Predictive Power

• Every 50 time steps, we build a model using data at $t_0$ to $t_i$, test prediction error at time steps $t_{i+1}$ to $t_{i+20}$ and calculate sum of squared error
• This shows us how accurate our model is over time
• As a baseline, we can assume a model that only predicts the mean value of the training set every time
• ARMA model alone doesn’t forecast large changes well

Future Directions

• Modeling both normal gameplay flow and important events
  • Capturing large, frequent changes in character attributes in a forecasting model
  • Predicting major events (e.g. character deaths, large battles, etc.)
  • Incorporating multivariate models using other character attributes
  • Modeling of multiplayer interactions
  • Clustering similar player’s behavior to understand types of players
  • Analyze interaction between groups of players between teams and within teams
• Quantifying uncertainty of both attribute forecasts and event prediction

References