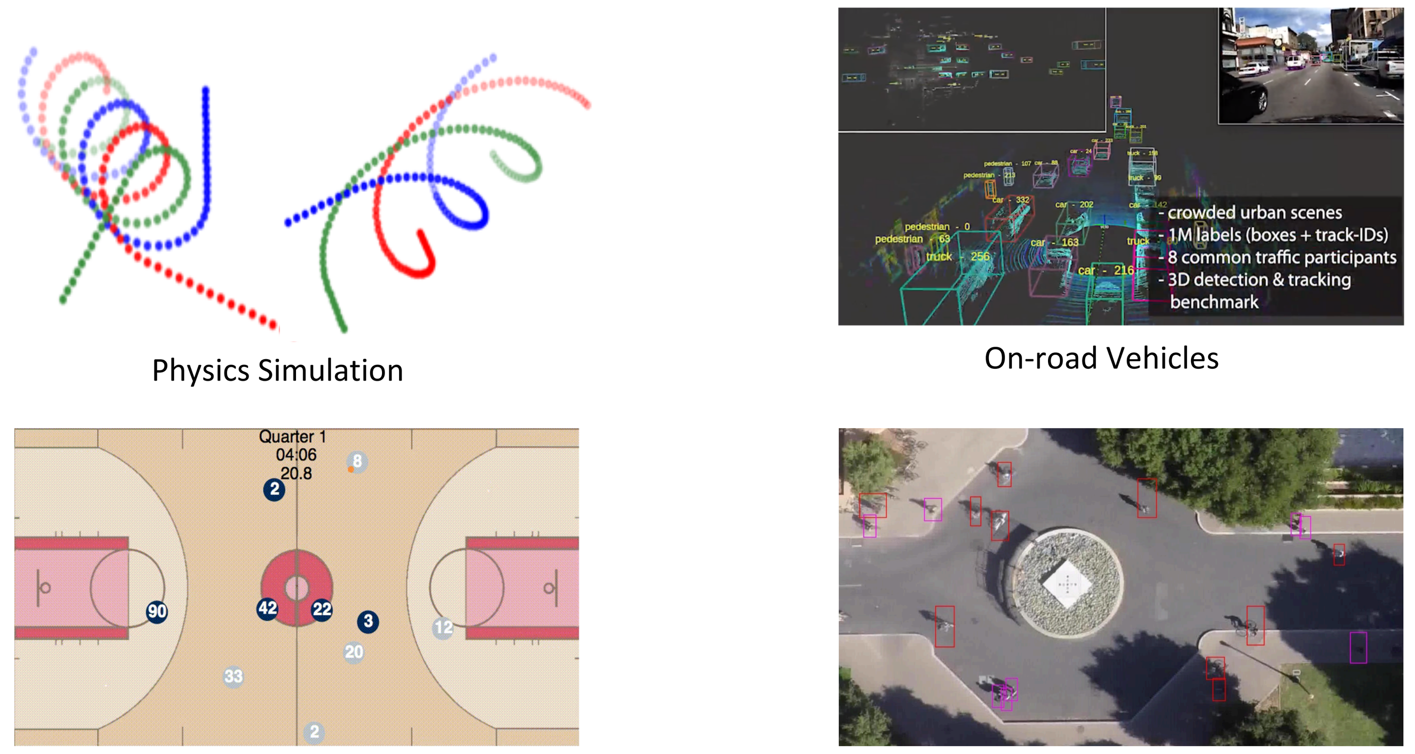
EvolveGraph: Multi-Agent Trajectory Prediction with Dynamic Relational Reasoning

Jiachen Li*, Fan Yang*, Masayoshi Tomizuka, Chiho Choi

Honda Research Institute US

Background and Goals



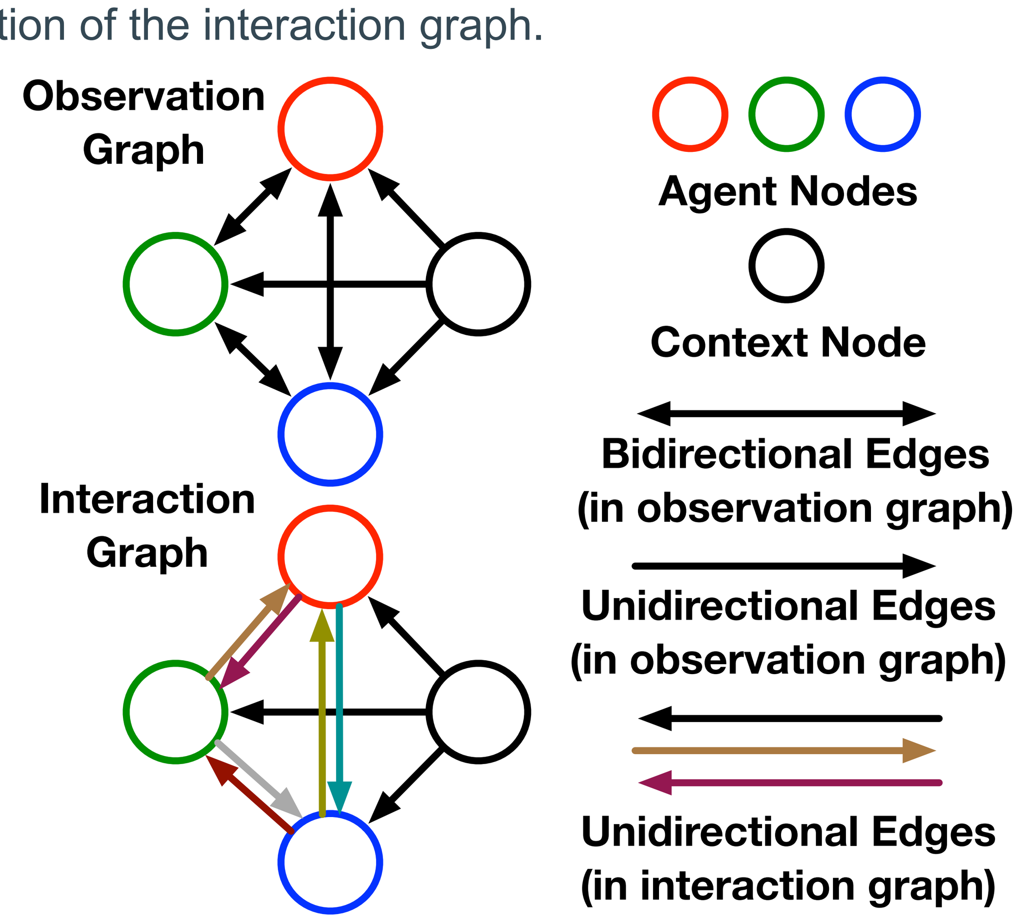
Accurate multi-agent trajectory prediction is critical in many real-world applications, where a group of entities interact with each other, giving rise to complicated behavior patterns of both individuals and the whole multi-agent system.

In this work, we address the problem of 1) extracting the underlying interaction patterns with a latent graph structure, which can handle different types of agents in a unified way, 2) capturing the dynamics of interaction graph evolution for dynamic relational reasoning, 3) predicting future trajectories (state sequences) based on the historical observations and the latent interaction graph, and 4) capturing the multi-modality of future system behaviors.

Observation Graph and Interaction Graph

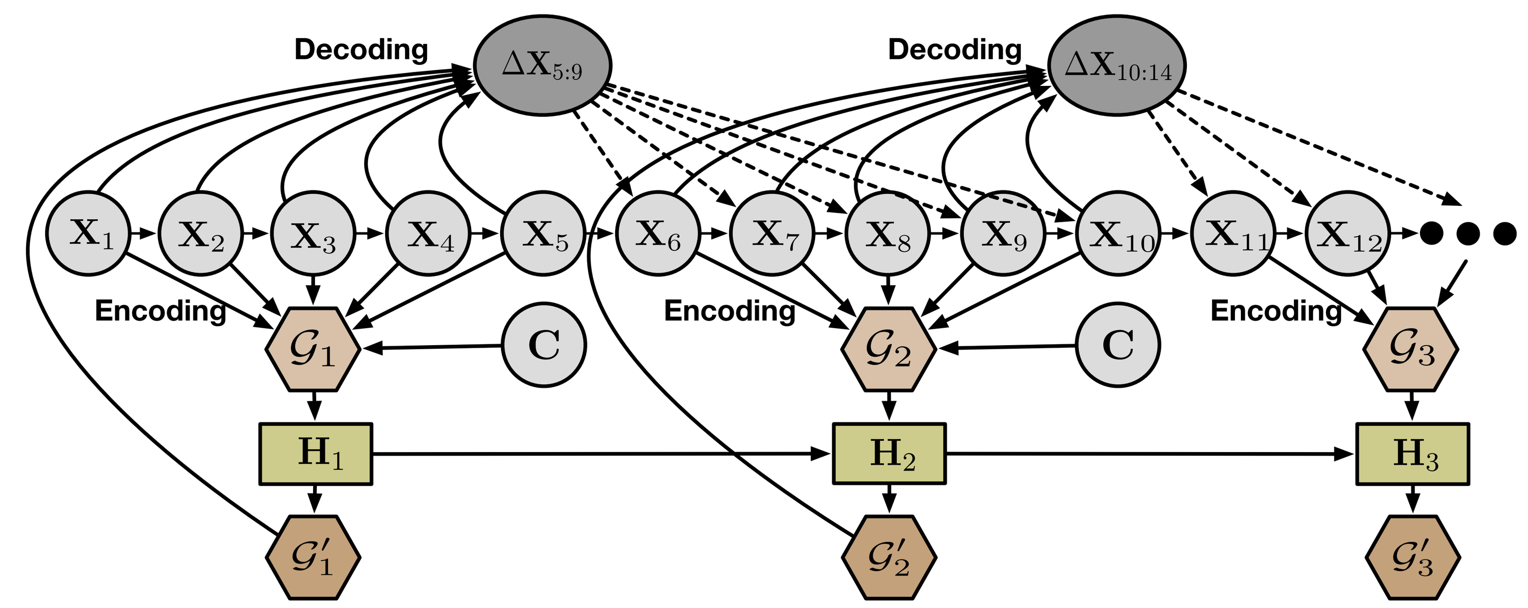
Observation Graph: The observation graph consists of N agent nodes and one context node. Agent nodes are bidirectionally connected to each other, and the context node only has outgoing edges to each agent node. Each agent node has two types of attributes: *self-attribute* and *social-attribute*.

Interaction Graph: The interaction graph represents interaction patterns with a distribution of edge types for each edge, which is built on top of the observation graph. The interaction graph is obtained by the encoding process. A recurrent unit is employed to enable the evolution of the interaction graph.



Project Website

EvolveGraph

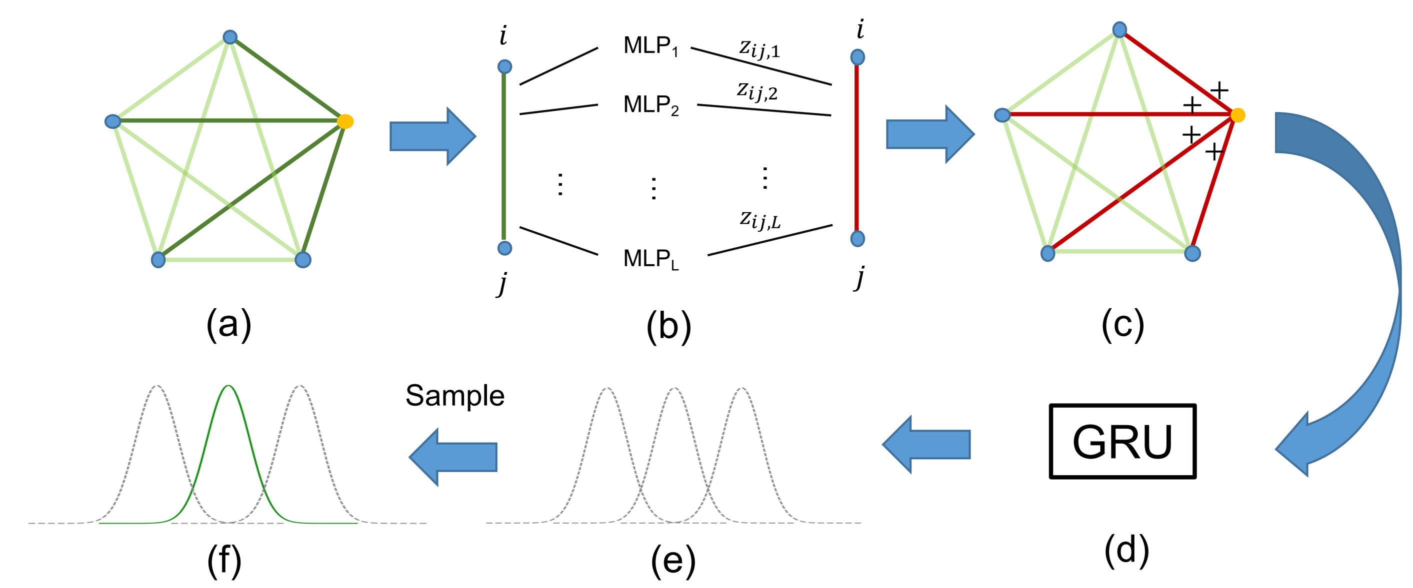


Encoding

The goal of the encoding process is to infer a latent interaction graph from the observation graph, which is essentially a multi-class edge classification task.

Decoding

A recurrent decoding process is applied to the interaction graph and observation graph to approximate the distribution of future trajectories.



Dynamic Interaction Graph Learning

In many applications, the interaction patterns recognized from the past time steps are likely not static in the future. Moreover, many interaction systems have multi-modal properties in its nature. Different modalities afterwards are likely to result in different interaction patterns. Therefore, we designed a dynamic evolving process of the interaction patterns.

The encoding process is repeated every τ (re-encoding gap) time steps to obtain the latent interaction graph based on the latest observation graph.

A recurrent unit (GRU) is utilized to maintain and propagate the history information, as well as adjust the prior interaction graphs.

Uncertainty and Multi-Modality

In our framework, the uncertainty and multi-modality mainly come from three aspects: a) In the decoding process, we output Gaussian mixture distributions indicating that there are several possible modalities at the next step. b) Different sampled trajectories will lead to different interaction graph evolution. The evolution of interaction graphs contributes to the multi-modality of future behaviors. c) Directly training such a model, however, tends to collapse to a single mode. Therefore, we employ an effective mechanism to mitigate the mode collapse issue and encourage multi-modality.

Experiments and Results

Particle Physics System

Multiple particles are initially linked and move together. The links disappear as if a certain criterion on particle state is satisfied and all the particles move independently thereafter.

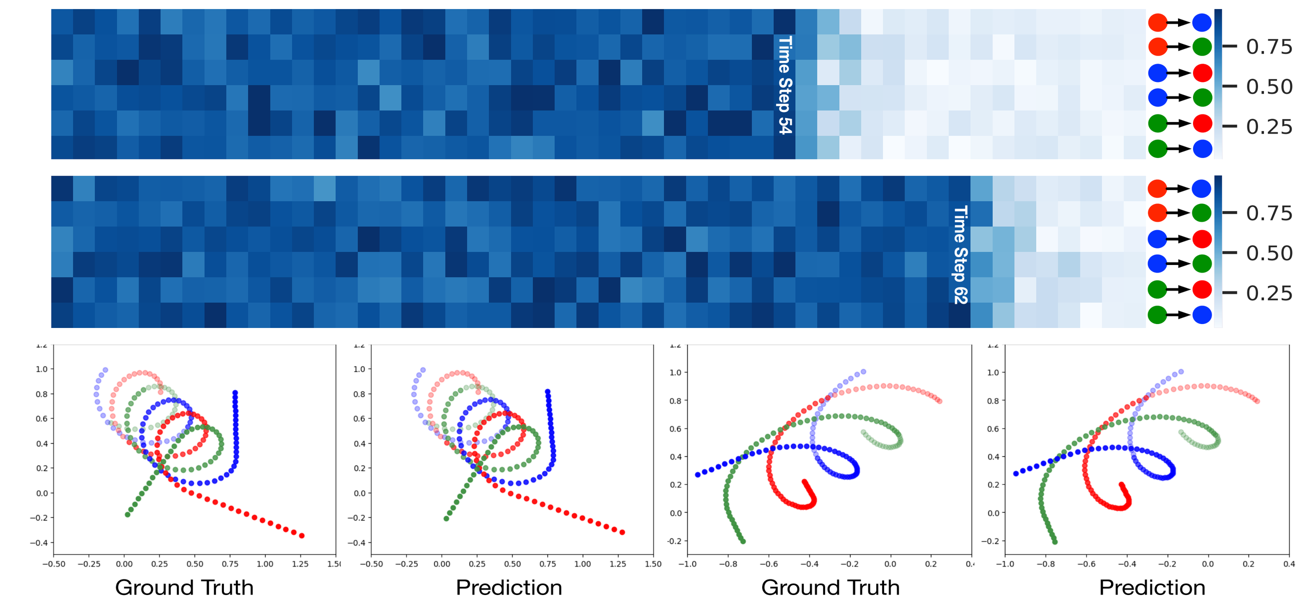


Table 1: Comparison of Accuracy (Mean ± Std in %) of Interaction (Edge Type) Recognition.

	Corr. (LSTM)	NRI (dynamic)	**EvolveGraph** (static)	**EvolveGraph** (RNN re-encoding)	**EvolveGraph** (dynamic)	Supervised
No Change	63.2±0.9	91.3±0.3	**95.6**±0.2	91.4±0.3	93.8±1.1	98.1±0.4
Change	—	71.5±3.1	64.1±0.8	75.2±1.4	**82.3**±3.2	94.3±1.5

Traffic Scenarios

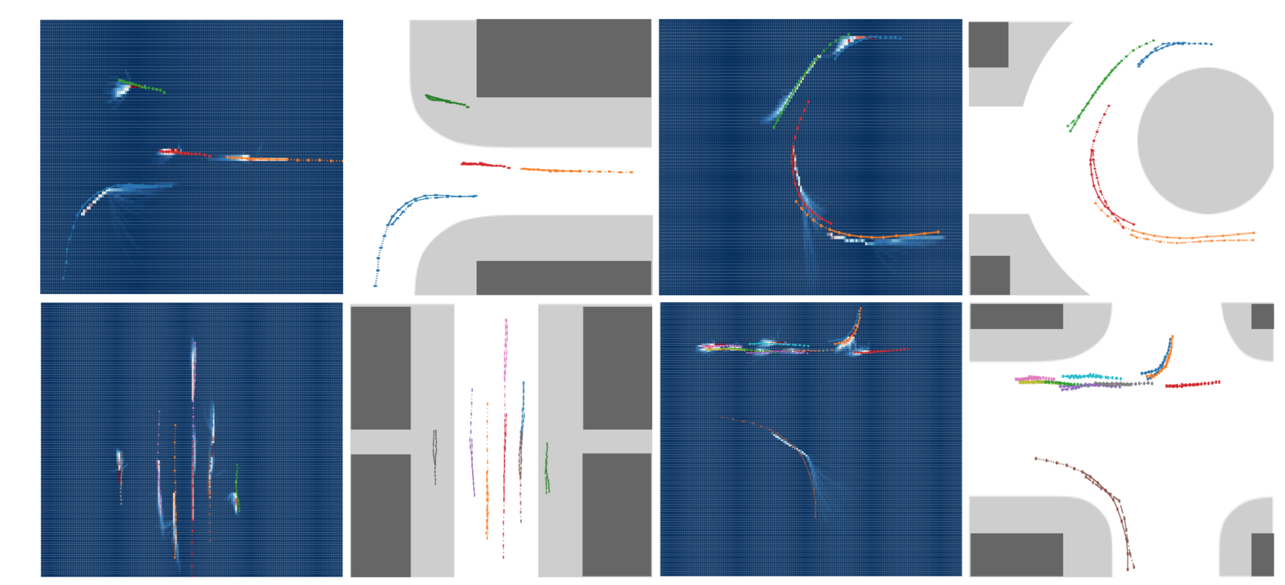


Table 2: minADE_{20} / minFDE_{20} (Meters) of Trajectory Prediction (H3D dataset).

Time	Baseline Methods							**EvolveGraph (Ours)**				
	STGAT	Social-Attention	Social-STGCNN	Social-GAN	Gated-RN	Trajectron++	NRI (dynamic)	SG (same node type)	SG	RNN re-encoding	DG (single stage)	DG (double stage)
1.0s	0.24 / 0.33	0.29 / 0.45	0.23 / 0.32	0.27 / 0.37	**0.18** / 0.32	0.21 / 0.34	0.24 / 0.30	0.28 / 0.37	0.27 / 0.35	0.25 / 0.32	0.24 / 0.31	0.19 / **0.25**
2.0s	0.34 / 0.48	0.53 / 0.96	0.36 / 0.52	0.45 / 0.77	0.32 / 0.64	0.33 / 0.62	0.32 / 0.60	0.40 / 0.58	0.38 / 0.55	0.35 / 0.51	0.33 / 0.46	**0.31 / 0.44**
3.0s	0.46 / 0.77	0.87 / 1.62	0.49 / 0.89	0.68 / 1.29	0.49 / 1.03	0.46 / 0.93	0.48 / 0.94	0.51 / 0.80	0.48 / 0.76	0.44 / 0.70	0.40 / 0.60	**0.39 / 0.58**
4.0s	0.60 / 1.18	1.21 / 2.56	0.73 / 1.49	0.94 / 1.91	0.69 / 1.56	0.71 / 1.63	0.73 / 1.56	0.64 / 1.21	0.61 / 1.14	0.57 / 1.07	0.50 / 0.90	**0.48 / 0.86**

Sports Players

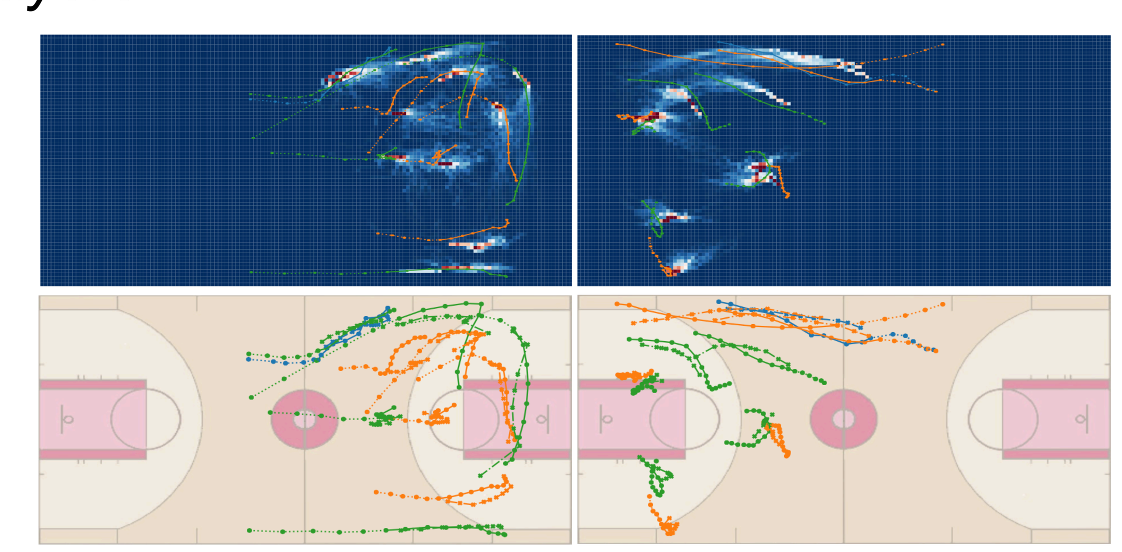


Table 3: minADE_{20} / minFDE_{20} (Meters) of Trajectory Prediction (NBA dataset).

Time	Baseline Methods							**EvolveGraph (Ours)**				
	STGAT	Social-STGCNN	Social-Attention	Social-LSTM	Social-GAN	Trajectron++	NRI (dynamic)	SG (same node type)	SG	RNN re-encoding	DG (single stage)	DG (double stage)
1.0s	0.42 / 0.71	0.46 / 0.76	0.87 / 1.36	0.92 / 1.34	0.82 / 1.25	0.55 / 0.90	0.60 / 0.87	0.70 / 1.09	0.59 / 0.92	0.58 / 0.89	0.48 / 0.76	**0.31 / 0.52**
2.0s	0.91 / 1.39	0.90 / 1.43	1.58 / 2.51	1.64 / 2.74	1.52 / 2.45	0.99 / 1.58	1.02 / 1.71	1.51 / 2.38	1.38 / 2.12	1.09 / 1.88	0.84 / 1.43	**0.74 / 1.10**
3.0s	1.62 / 2.87	1.59 / 2.67	2.78 / 4.66	2.93 / 5.03	2.63 / 4.51	1.89 / 3.32	1.83 / 3.15	2.10 / 3.53	1.88 / 3.23	1.77 / 2.87	1.43 / 2.55	**1.28 / 2.07**
4.0s	2.47 / 3.86	2.35 / 3.71	3.76 / 6.64	4.00 / 7.12	3.60 / 6.24	2.62 / 4.70	2.48 / 4.30	2.83 / 4.85	2.52 / 4.57	2.39 / 3.89	2.08 / 3.74	**1.83 / 3.16**