TRANS System: Fundamental First Principles

The Core Thesis

"Sleeper stocks are coiled springs - the sequence generation captures both the coil shape (temporal) and the spring tension (context)."

First Principle: The Sleeper Hypothesis

The system is built on a single market microstructure insight:

Illiquid micro/small-cap stocks that consolidate in a tight range with drying volume are accumulating "potential energy" that releases explosively on breakout.

This is NOT momentum trading. This is NOT trend following. This is detecting supply exhaustion in forgotten corners of the market.

What Makes a "Sleeper"?

A sleeper has three characteristics the system must detect:

Characteristic	What It Means	How System Detects
Dormancy	Stock is forgotten, volume dri..	dormancy_shock, vol_dryup_rati..
Coil	Price compressed in tight rang..	bbw_20, boundary slopes, coil_..
Sideways Regime	Not momentum, not crashing	trend_position (0.6-1.3 x SMA_..

The Two-Branch Architecture

The sequence generation creates two complementary views of each pattern:

Branch A: Temporal Sequences (10 features x 20 timesteps)

"The Movie" - How did the coil form over time?

```
Day 1  -> Day 20 (sliding window)

   +-- OHLC [0-3]: Price action relativized to day 0

   |      - Shows compression/expansion over time

   +-- Volume [4]: Wake-up detection

   |      - Normalized to 6-month dormant baseline, NOT day-0

   |

   |      - BBW_20: Bollinger width (tight = coiled)

   |      - Volume Ratio: Recent vs 20d avg

   +-- Boundary Slopes [8-9]: Triangle geometry

          - lower_slope: Rolling regression of lows

          - Rectangle: flat (both ~ 0)
```

Branch B: Context Features (14 static features)

"The Snapshot" - What is the potential energy at detection moment?

```
At pattern end date (single vector):

   +-- Market Structure [0-6]:

   |      - trend_position: Where in macro trend

   |      - log_float, log_dollar_volume: Tradability

   +-- Deep Dormancy [7-8]:

   |        -> "Is current activity the highest in a YEAR?"

   |        -> "How exhausted is supply?"

   +-- Coil State [9-13]:

          - distance_to_danger: Proximity to breakdown

          - coil_intensity: Combined quality score
```

The Filtering Philosophy

"Garbage in, garbage out" - Before any ML, ruthlessly filter to tradeable sleepers:

Layer 1: Physics Filter (Invalid -> DROP)

Filter	Rejects	Why
Market Cap	Large/Mega caps	No explosive moves
Width	< 2% patterns	Untradeable (spread eats profi..
Dollar Volume	< $50k/day	Can't get filled
Zombie Health	Ghost trades, data gaps	Data errors, not real patterns
Sideways Regime	< 0.6 or > 1.3 x SMA_200	Crashing or momentum (not slee..

Layer 2: NMS (Overlap -> DEDUPE)

- One consolidation event can trigger multiple detections

Layer 3: Heartbeat Filter (Erratic -> MARK as Noise)

- High volume CV = erratic trading pattern

The Labeling Ground Truth

After 100 days, label what ACTUALLY happened:

```
Entry: Upper boundary of consolidation box

Target: Entry + 5R

Class 0 (Danger): Hit -2R stop first     -> Strategic Value: -2.0

Class 2 (Target): Hit +5R target first   -> Strategic Value: +5.0
```

Key insight: Labels are PURE ground truth. A breakout is a breakout. Width/tradeability filtering happens at inference, not labeling.

The Model's Job

Given the temporal sequence and context snapshot, predict:

```
EV = P(Danger) x (-2.0) + P(Noise) x (-0.1) + P(Target) x (+5.0)
```

Not "will it break out?" but "what's the expected value of trading this pattern?"

Why This Architecture Works for Illiquid Markets

Challenge	Solution
Zero volume = signal, not miss..	No Gaussian noise augmentation..
Wide spreads = noisy prices	Split attention (price group v..
Breakout can happen any day	RoPE + Window Jittering (posit..
Arbitrary pattern start	Volume normalized to 6-month m..
Float matters for move potenti..	Context-conditioned LSTM (h0/c..

Summary: The Fundamental Formula

```
Sleeper Detection = Temporal Coil Shape + Static Potential Energy + Strict Fi...
```
```
Where:
```
```
- Potential Energy = 14 context features (GRN -> LSTM conditioning)
```

The system doesn't predict price. It predicts which consolidating sleepers have accumulated enough potential energy to be worth the risk.

Feature Summary Tables

10 Temporal Features (per timestep)

Index	Feature	Description
0	open	(open_t / close_0) - 1
1	high	(high_t / close_0) - 1
2	low	(low_t / close_0) - 1
3	close	(close_t / close_0) - 1
4	volume	log(volume_t / vol_6m_median)
5	bbw_20	Bollinger Band Width (20-perio..
6	adx	Average Directional Index
7	volume_ratio_20	Volume / 20-day average
8	upper_slope	Rolling regression slope of hi..
9	lower_slope	Rolling regression slope of lo..

14 Context Features (static at pattern end)

Index	Feature	Description
0	float_turnover	Sum(Vol_60D) x Price / Market_..
1	trend_position	Close / SMA_200
2	base_duration	log-normalized pattern duratio..
3	relative_volume	Vol_20D / Vol_60D
4	distance_to_high	(52W_High - Close) / 52W_High
5	log_float	log10(shares_outstanding)
6	log_dollar_volume	log10(avg_daily_dollar_volume)
7	dormancy_shock	log10(vol_20d / vol_252d)
8	vol_dryup_ratio	vol_20d / vol_100d
9	price_position_at_end	Position in box (0=lower, 1=up..
10	distance_to_danger	Distance from danger zone
11	bbw_slope_5d	BBW change over 5 days
12	vol_trend_5d	Recent volume vs 20d avg
13	coil_intensity	Combined coil quality score

Document generated: January 2026