

HAI Dataset 설명

- HAI (HIL-based Augmented ICS) Dataset
 - Hardware-In-the-Loop (HIL) 시뮬레이터를 통하여 보강된 실제 산업 제어 시스템 (Industrial Control System, ICS) 테스트베드에서 수집된 데이터
 - 테스트베드에서는 증기-터빈 발전 (steam-turbine power generation) 과 양수-저장 수력 발전 (pumped-storage hydropower generation) 을 모방



HAI DATASET

HIL-BASED AUGMENTED ICS (HAI) DATASET WAS COLLECTED FROM A REALISTIC ICS TESTBED AUGMENTED WITH A HARDWARE-IN-THE-LOOP SIMULATOR THAT EMULATES STEAM-TURBINE POWER GENERATION AND PUMPED-STORAGE HYDROPOWER GENERATION

HAI Dataset 설명

- HAI (HIL-based Augmented ICS) Dataset
 - 국가보안기술연구소가 공개한 보안 데이터셋으로 산업제어시스템 보안기술 개발에 적합
 - 수작업에 의존한 기존 데이터셋의 데이터 라벨링 한계점을 극복해 데이터 신뢰성을 보장하고, 공격 수준별 데이터 생성이 가능하며 정밀한 성능평가가 가능함
 - 현재 세계적으로 가장 많이 사용되고 있는 싱가포르 SUTD 대학의 SWaT 데이터셋과 비슷한 크기



HAI DATASET

HIL-BASED AUGMENTED ICS (HAI) DATASET WAS COLLECTED FROM A REALISTIC ICS TESTBED AUGMENTED WITH A HARDWARE-IN-THE-LOOP SIMULATOR THAT EMULATES STEAM-TURBINE POWER GENERATION AND PUMPED-STORAGE HYDROPOWER GENERATION

HAI Dataset 설명

- HAI (HIL-based Augmented ICS) Dataset
 - HAI 데이터셋은 v1과 v2가 있음
 - v1은 59개의 특징들로 이루어져 있으며, v2는 78개의 특징들로 이루어져 있음
 - v1에서는 데이터 설명이 있어서 각 특징이 의미하는 바, 데이터들 간의 관계, 단위 등을 알 수 있지만, v2는 알 수 없음



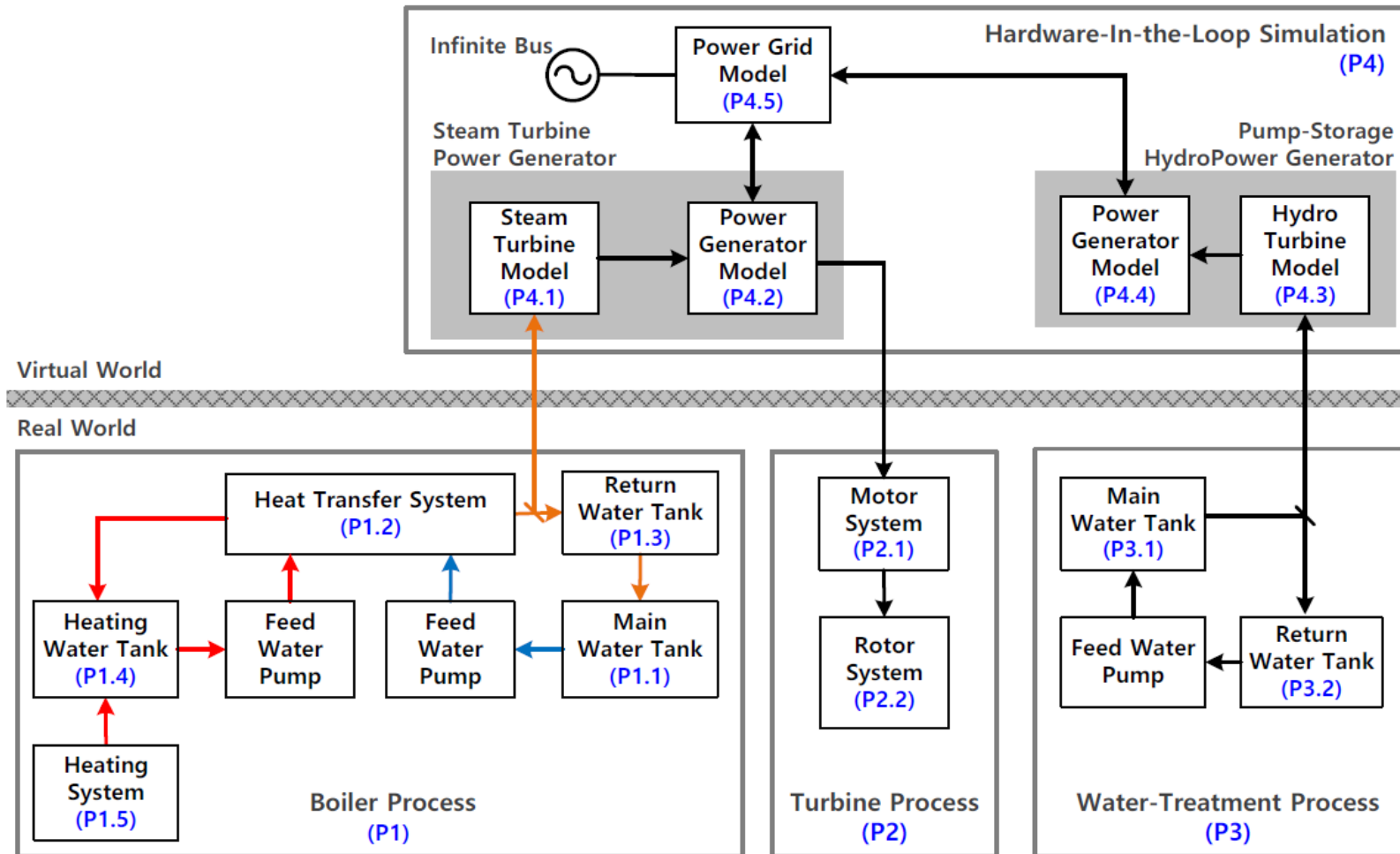
HAI DATASET

HIL-BASED AUGMENTED ICS (HAI) DATASET WAS COLLECTED FROM A REALISTIC ICS TESTBED AUGMENTED WITH A HARDWARE-IN-THE-LOOP SIMULATOR THAT EMULATES STEAM-TURBINE POWER GENERATION AND PUMPED-STORAGE HYDROPOWER GENERATION

HAI Dataset 설명

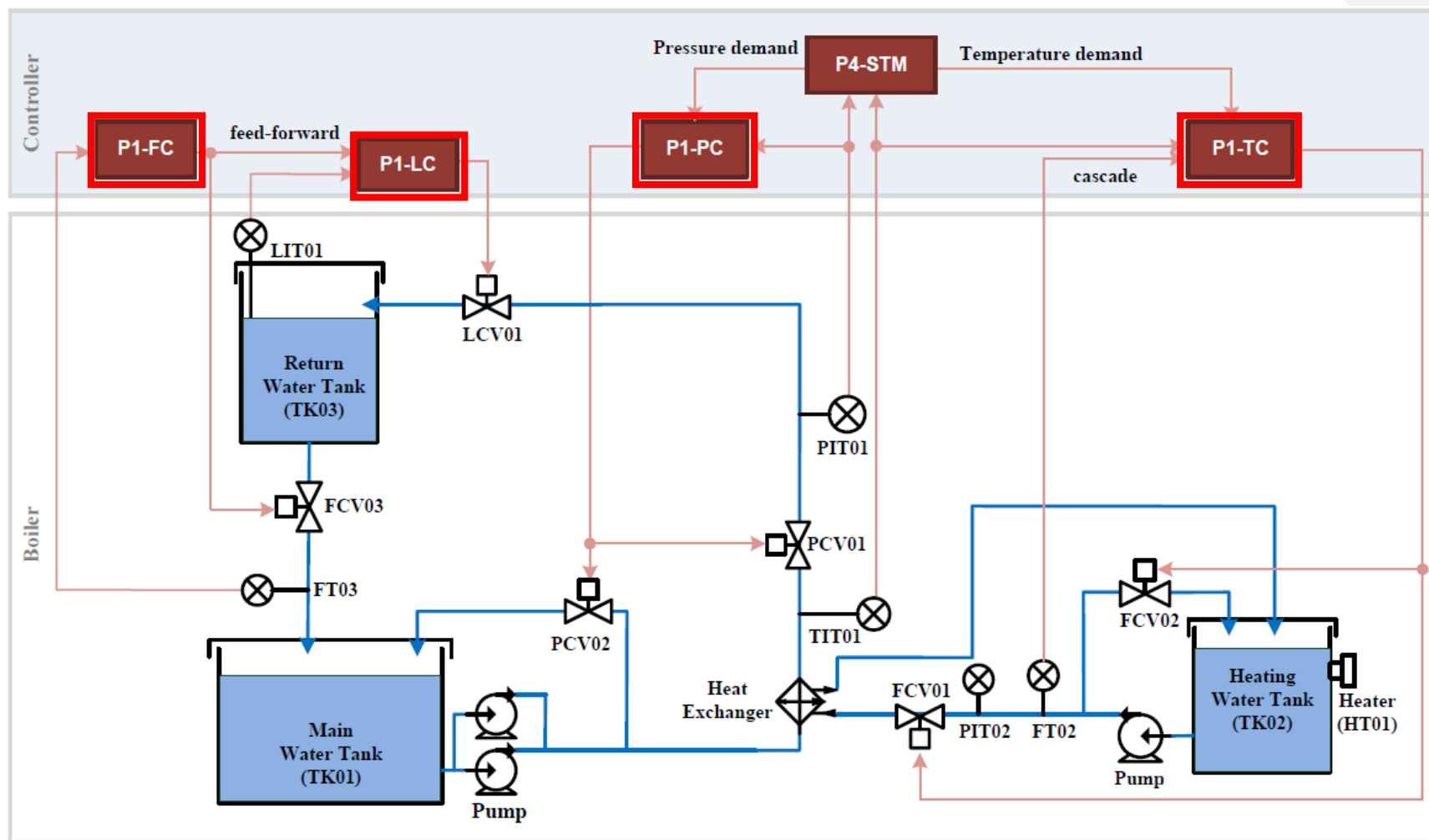
- HAI Dataset

- 4가지 공정(보일러, 터빈, 정수 처리, HIL 시뮬레이션)을 포함함



HAI Dataset 설명

■ 보일러 조절



HAI Dataset 설명

- 보일러 조절
 - 관련 데이터

No	Point Name	Range	Unit	Description
1	P1_B2004	0 ~ 10	Bar	Heat-exchanger outlet pressure setpoint
2	P1_B2016	0 ~ 10	bar	Pressure demand to follow P1_B2004 and electrical load from steam turbine model
3	P1_B3004	0 ~ 720	mm	Water level setpoint in return water tank
4	P1_B3005	0 ~ 2,500	L/H	Water outflow rate setpoint from return water tank
5	P1_B4002	0 ~ 100	°C	Heat-exchanger outlet temperature setpoint
6	P1_B4005	0 ~ 100	-	Temperature cascade control (On: 1, Off: 0)
7	P1_B400B	0 ~ 3,000	L/H	Water outflow rate setpoint from heating water tank
8	P1_B4022	0 ~ 100	°C	Temperature demand to follow P1. B4005 and electrical load from steam-turbine model
9	P1_FCV01D	0 ~ 100	%	Position command for FCV01 valve
10	P1_FCV01Z	0 ~ 100	%	Current position of FCV01 valve
11	P1_FCV02D	0 ~ 100	%	Position command for FCV02 valve
12	P1_FCV02Z	0 ~ 100	%	Current position of FCV02 valve
13	P1_FCV03D	0 ~ 100	%	Position command for FCV03 valve
14	P1_FCV03Z	0 ~ 100	%	Current position of FCV03 valve
15	P1_FT01	0 ~ 2500	mmH ₂ O	Digital value of FT01 flow transmitter

HAI Dataset 설명

- 보일러 조절
 - 관련 데이터

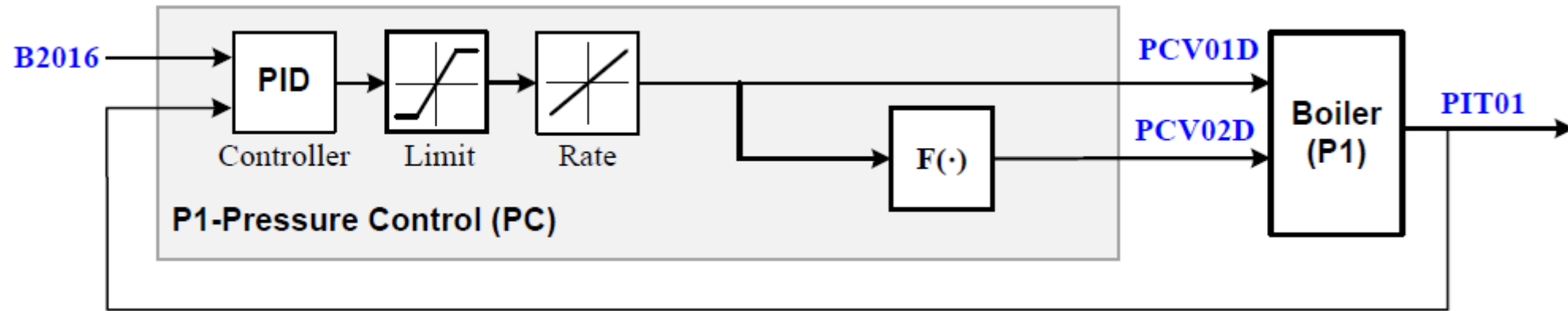
No	Point Name	Range	Unit	Description
16	P1_FT01Z	0 ~ 3190	L/H	Water inflow rate into return water tank
17	P1_FT02	0 ~ 2500	mmH ₂ O	Digital value of FT02 flow transmitter
18	P1_FT02Z	0 ~ 3190	L/H	Conversion from P1_FT02 to outflow rate at heating water tank
19	P1_FT03	0 ~ 2500	mmH ₂ O	Digital value of FT03 flow transmitter
20	P1_FT03Z	0 ~ 3190	L/H	Conversion from P1_FT03 to outflow rate at return water tank
21	P1_LCV01D	0 ~ 100	%	Position command for LCV01 valve
22	P1_LCV01Z	0 ~ 100	%	Current position of LCV01 valve
23	P1_LIT01	0 ~ 720		Water level of return water tank
24	P1_PCV01D	0 ~ 100	%	Position command for PCV01 valve
25	P1_PCV01Z	0 ~ 100	%	Current position of PCV01 valve
26	P1_PCV02D	0 ~ 100	%	Position command for PCV2 valve
27	P1_PCV02Z	0 ~ 100	%	Current position of PCV02 valve
28	P1_PIT01	0 ~ 10	bar	Heat-exchanger outlet pressure
29	P1_PIT02	0 ~ 10	bar	Water supply pressure of heating water pump
30	P1_TIT01	-10 ~ 100	°C	Heat-exchanger outlet temperature
31	P1_TIT02	-10 ~ 100	°C	Temperature of heating water tank

HAI Dataset 설명

■ 보일러 조절

➤ 압력 조절

- 두개의 압력 제어 밸브를 제어하여 설정에 따라 메인 물 탱크와 회수 물 탱크 사이의 압력을 유지하도록 제어

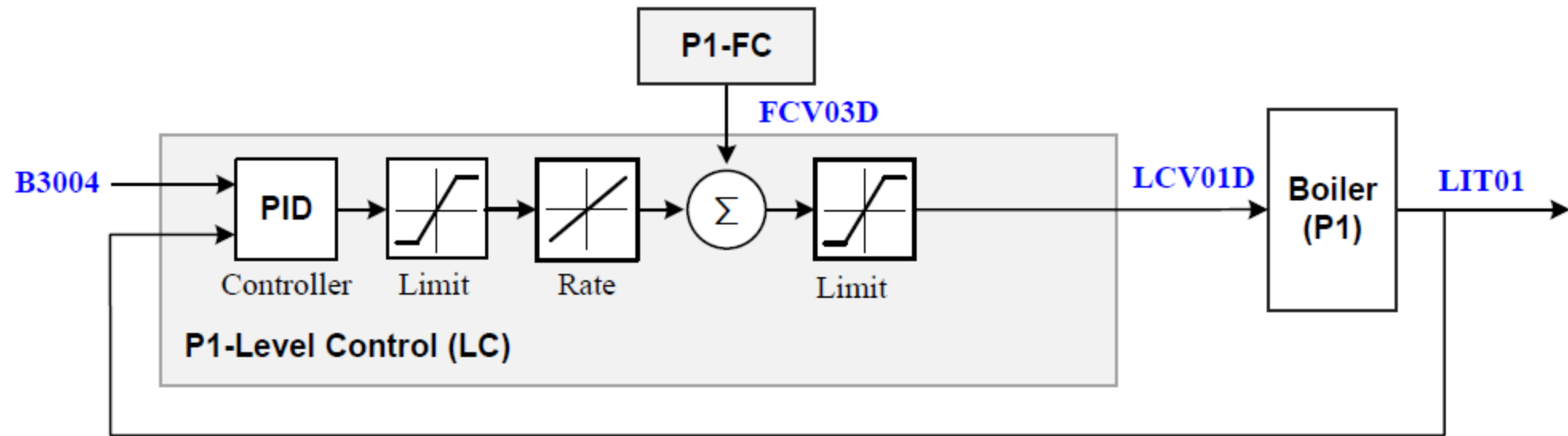


HAI Dataset 설명

- 보일러 조절

- 수위 조절

- 설정에 따라 회수 물 탱크 사이의 수위를 유지하기 위하여 밸브를 제어

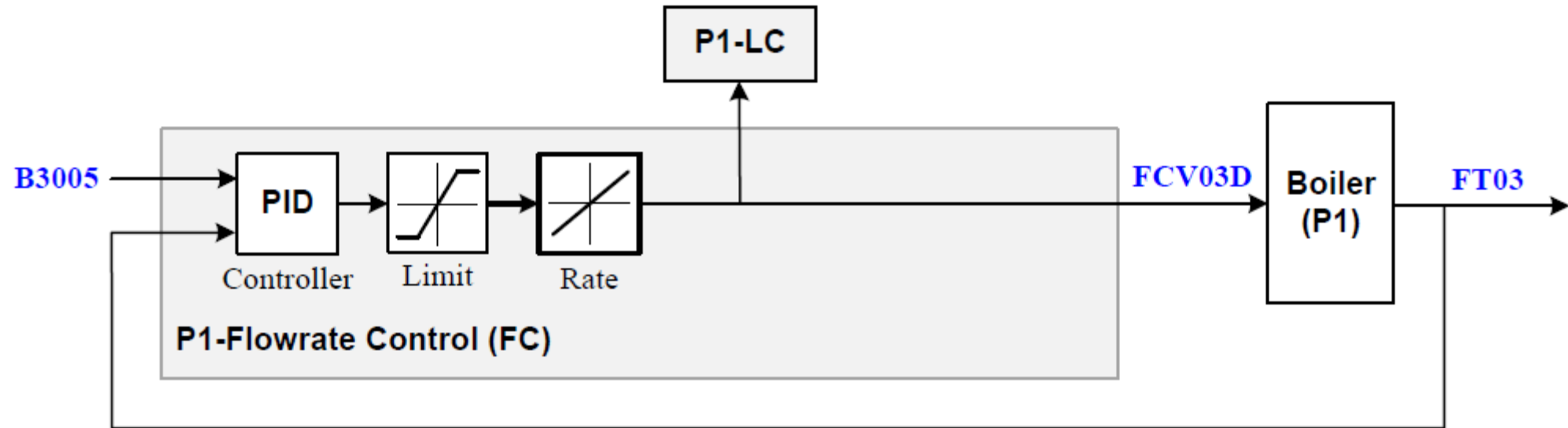


HAI Dataset 설명

- 보일러 조절

- 유속 조절

- 밸브를 제어하여 설정에 따라 회수 물탱크의 유출량을 유지하도록 제어

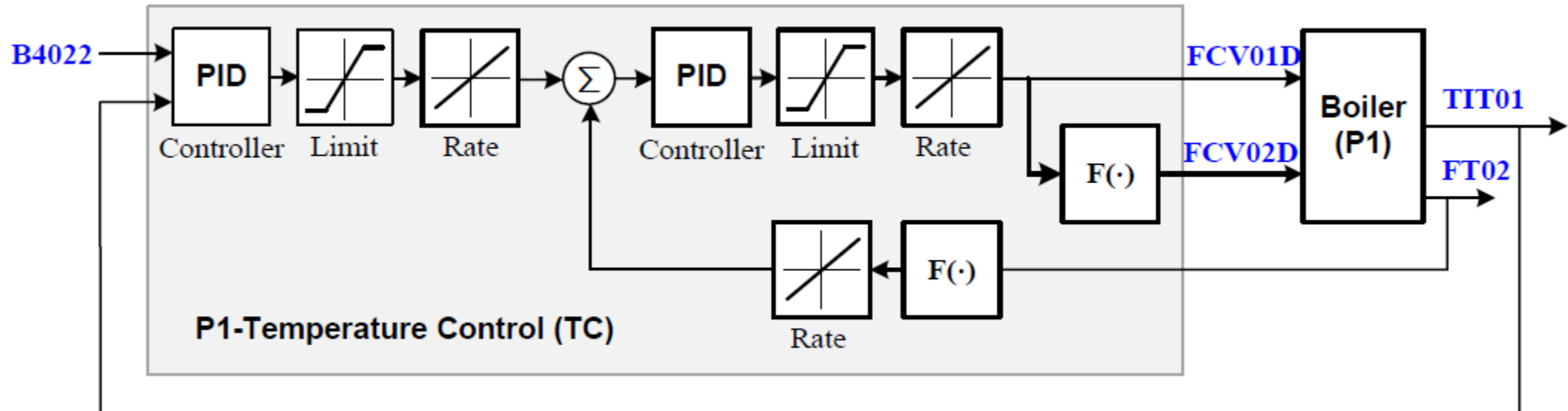


HAI Dataset 설명

■ 보일러 조절

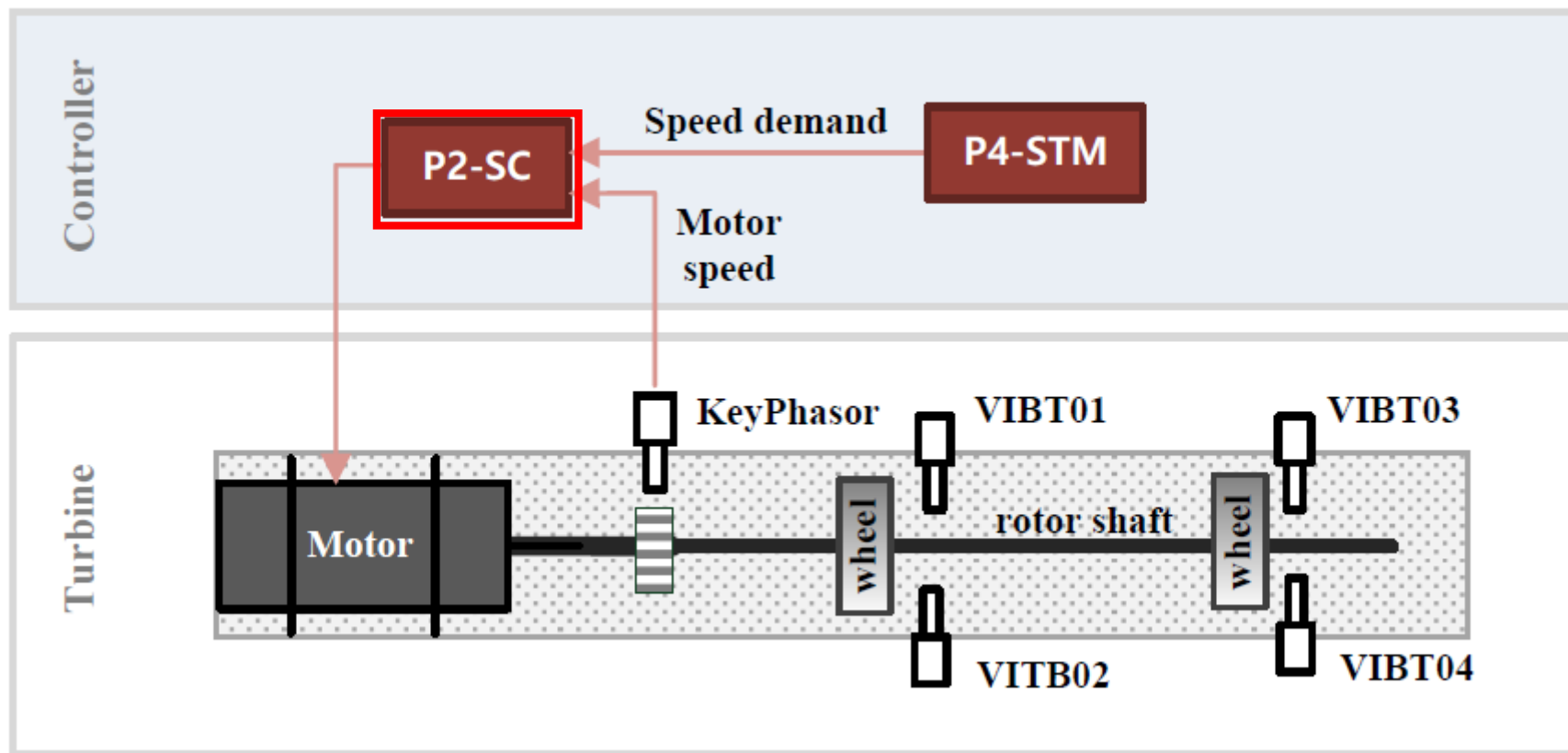
➤ 온도 조절

- 두개의 밸브를 제어하여 설정에 따라 회수 메인 용기의 온도를 유지



HAI Dataset 설명

- 터빈 조절



HAI Dataset 설명

- 터빈 조절
 - 관련 데이터

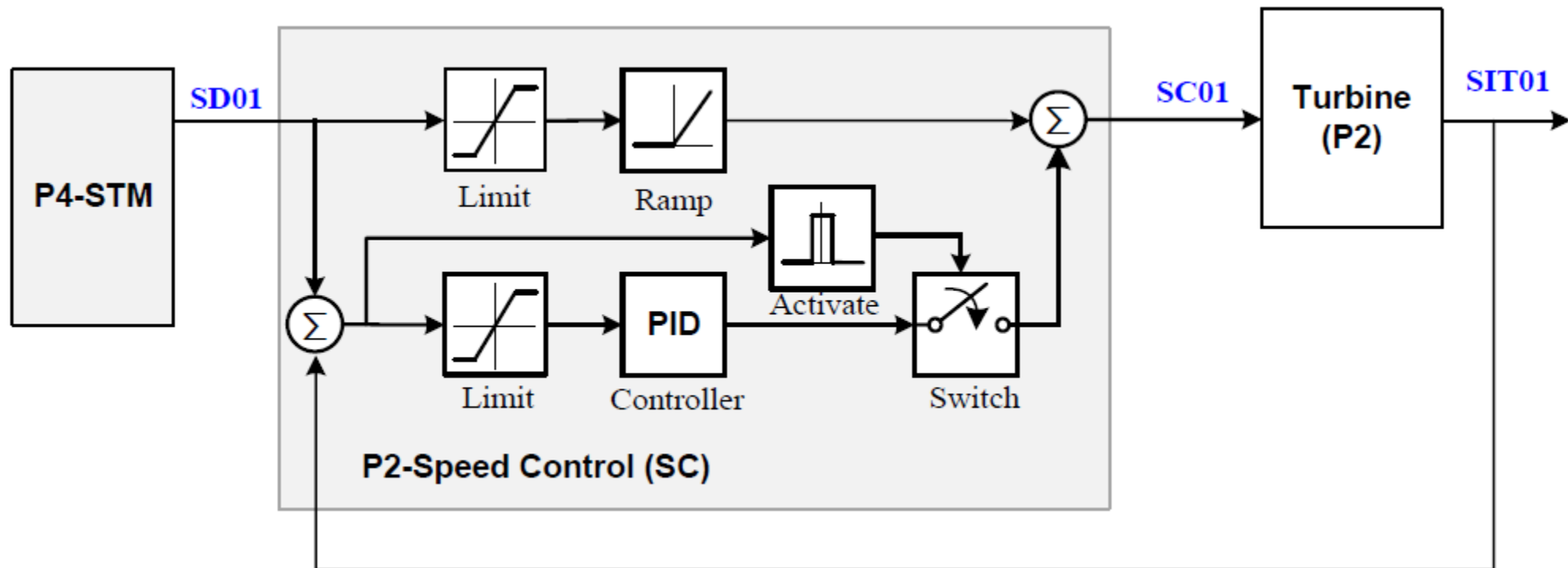
No	Point Name	Range	Unit	Description
32	P2_SIT01	0 ~ 3600	RPM	Current motor speed
33	P2_SD01	0 ~ 3600	RPM	User speed demand
34	P2_VT01	0 ~ 15	V	Phase lag signal of key phasor probe near motor
35	P2_VYT02	-10 ~ 10	μm	Shaft-vibration-related Y-axis displacement near the first mass wheel
36	P2_VXT02	-10 ~ 10	μm	Shaft-vibration-related X-axis displacement near the first mass wheel
37	P2_VYT03	-10 ~ 10	μm	Shaft-vibration-related Y-axis displacement near the second mass wheel
38	P2_VXT03	-10 ~ 10	μm	Shaft-vibration-related X-axis displacement near the second mass wheel
39	P2_24Vdc	0 ~ 30	V	DCS power supply
40	P2_Auto	0 or 1	-	System auto/manual mode
41	P2_Emg	0 or 1	-	Emergency-stop input
42	P2_On	0 or 1	-	System on/off input
43	P2_TripEx	0 or 1	-	Trip exit input

HAI Dataset 설명

■ 터빈 조절

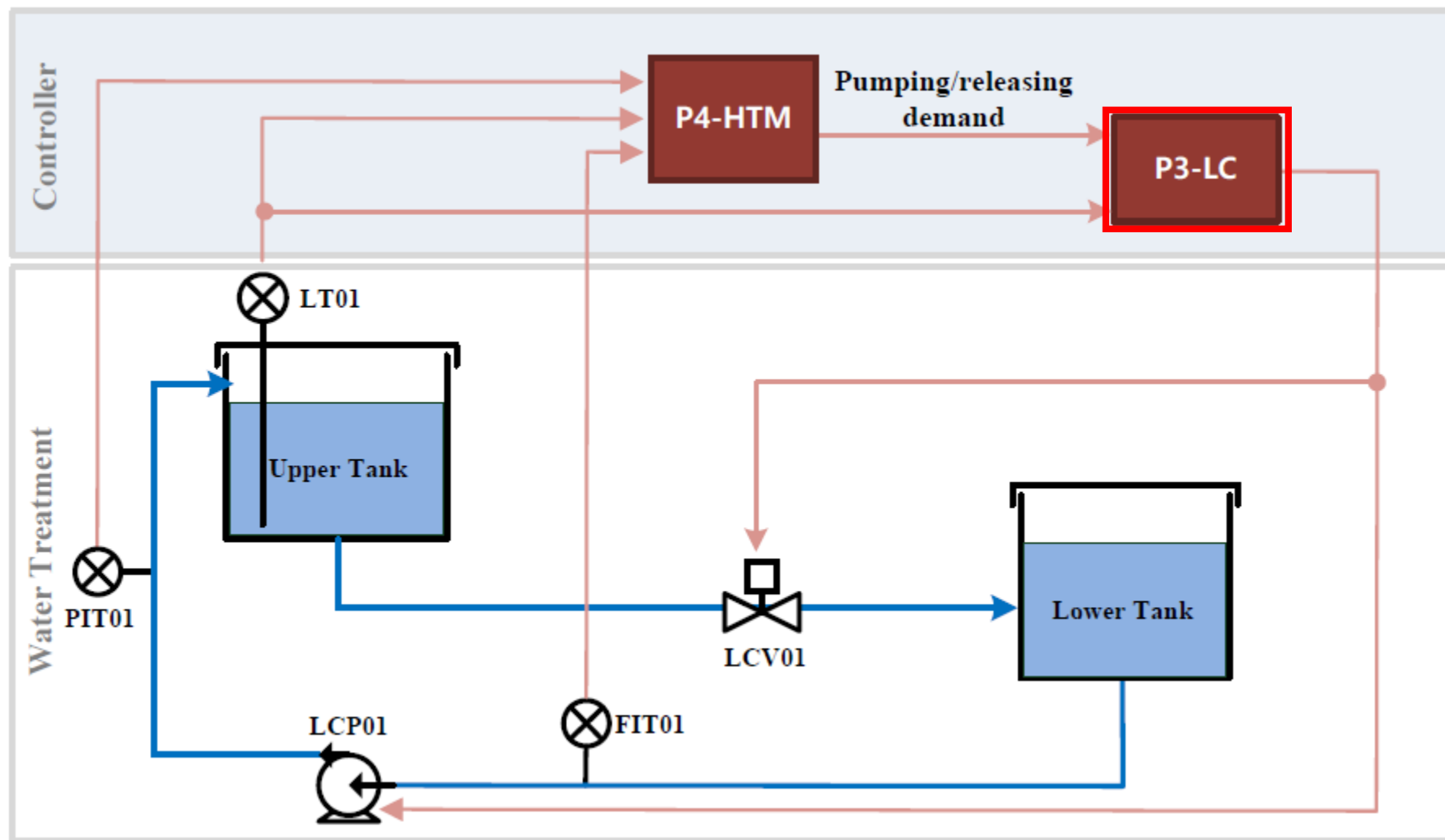
➤ 속도 조절

- 모터를 최소 제어 속도로 움직이게 만든 다음, PI 컨트롤러와의 결합 제어를 통해 모터의 속도를 설정한 값에 가까워지게 조절



HAI Dataset 설명

- 정수처리 조절



HAI Dataset 설명

- 정수처리 조절
 - 관련 데이터

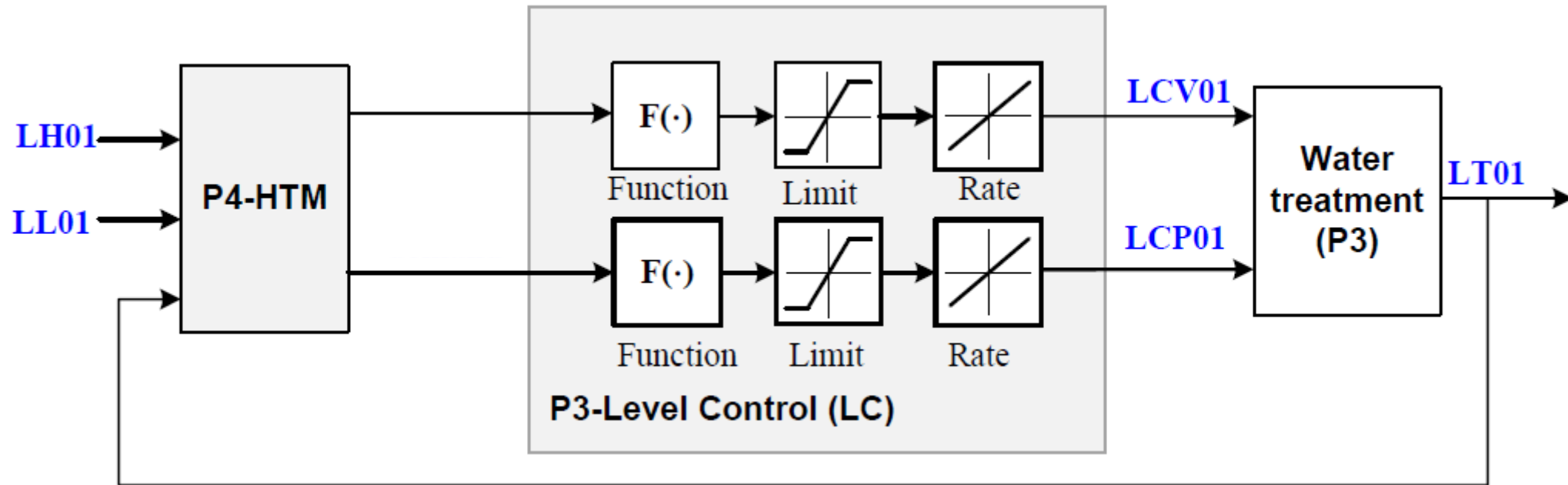
No	Point Name	Range	Unit	Description
44	P3_LT01	0 ~ 100	%	Water level in upper tank
45	P3_LH01	0 ~ 100	%	High water level setpoint
46	P3_LL01	0 ~ 100	%	Low water level setpoint
47	P3_LCP01D	0 ~ 27648	-	Speed command for feed water pump
48	P3_LCV01D	0 ~ 27648	-	Position command for LCV01 valve

HAI Dataset 설명

■ 정수처리 조절

➤ 수위 조절

- HIL 시뮬레이터의 배출 및 펌핑 요구량을 조정하여 밸브와 펌프를 조절



HAI Dataset 설명

- HIL 시뮬레이션 조절
 - 관련 데이터

No	Point Name	Range	Unit	Description
49	P4_LD	0 ~ 600	MW	Total electrical load demand
50	P4_ST_FD	-0.1 ~ 0.1	mHz	Frequency deviation of steam-turbine model
51	P4_ST_PO	0 ~ 500	MW	Output power of steam-turbine model
52	P4_ST_PT01	0 ~ 27648	-	Digital value of steam pressure in steam-turbine model
53	P4_ST_TT01	0 ~ 27648	-	Digital value of steam temperature in steam-turbine model
54	P4_ST_LD	0 ~ 500	MW	Electrical load demand for steam-turbine model
55	P4_ST_PS	0 ~ 500	MW	Scheduled power demand of steam-turbine model
56	P4_HT_FD	-0.1 ~ 0.1	mHz	Frequency deviation of hydropower-turbine model
57	P4_HT_PO	0 ~ 100	MW	Output power of hydropower-turbine model
58	P4_HT_LD	0 ~ 100	MW	Electrical load demand for steam-turbine model
59	P4_HT_PS	0 ~ 100	MW	Scheduled power demand of hydropower-turbine model

평가 지표

- 분류(Classification) 문제에 사용되는 평가 지표
 - 혼동 행렬(Confusion Matrix)을 기반으로 측정하는 정확도(Accuracy), 정밀도(Precision), 재현율(Recall), F1-Score 등이 있음
 - 혼동 행렬이란 모델의 분류 결과와 실제 정답을 행렬 형태로 아래와 같이 나타낸 것

		실제 정답	
		True	False
		True	False
분류 결과	True	True Positive	False Positive
	False	False Negative	True Negative

평가 지표

- 분류(Classification) 문제에 사용되는 평가 지표
 - 다양한 평가지표 중 정밀도와 재현율이 널리 사용됨
 - 정밀도 : 예측을 Positive로 한 대상 중 예측과 실제 값이 일치하는 비율
 - 재현율 : 실제 Positive인 대상 중에서 예측과 실제 값이 일치하는 비율

$$\text{Precision} = \frac{TP}{FP+TP}$$

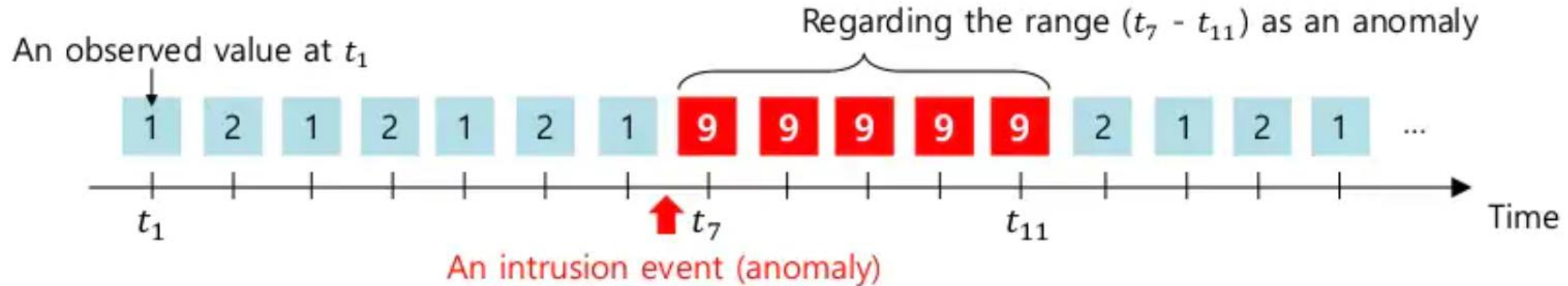
$$\text{Recall} = \frac{TP}{FN+TP}$$

		실제 정답	
		True	False
분류 결과	True	True Positive TP	False Positive FP
	False	False Negative FN	True Negative TN

		실제 정답	
		True	False
분류 결과	True	True Positive TP	False Positive FP
	False	False Negative FN	True Negative TN

평가 지표

- 시계열 데이터에서의 이상 탐지를 위한 평가지표
 - 시계열 데이터가 아닌 경우에는 단순히 데이터에 이상이 있는지 없는지만 판단하면 됨
 - 하지만, 시계열 데이터의 경우 범위도 굉장히 중요한 의미를 가지고 있음



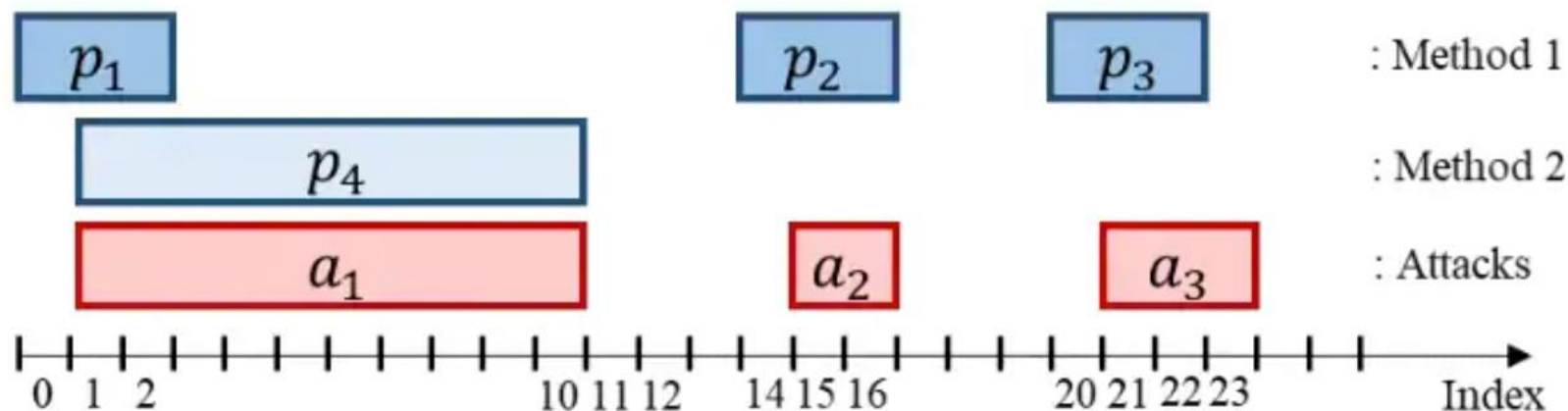
평가 지표

■ 시계열 데이터에서의 이상 탐지를 위한 평가지표

➤ 또한, 정밀도와 재현율은 다양한 원인으로 인하여 발생한 이상을 탐지하였는지를 평가하기에는 부적합

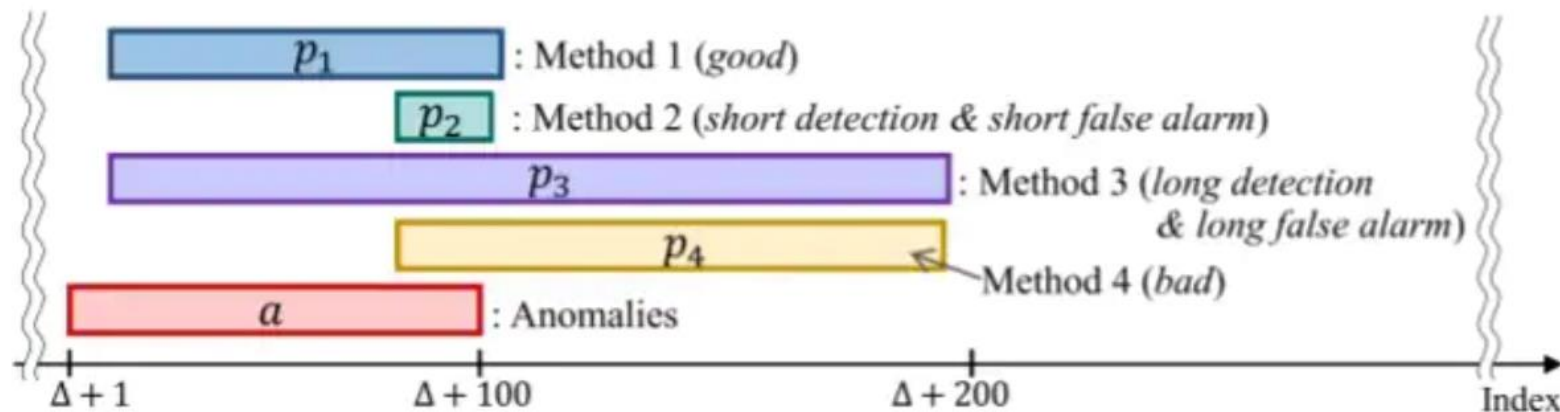
- 아래 그림에서 방법2는 a_1 만을 탐지하였으나 높은 점수를 받는 것을 확인할 수 있음

Method	Metric	
	Precision	Recall
1	0.67	0.40
2	1.00	0.67



평가 지표

- 시계열 데이터에서의 이상 탐지를 위한 평가지표
 - 데이터에 이상이 있는지에 대한 여부와 더불어 범위까지 측정할 수 있는 평가지표를 사용해야 함
 - TaPR은 위의 조건을 만족하는 평가를 제공
 - TaPR : Time-series aware Precision and Recall for anomaly detection
 - TaR과 TaP로 구성이 되어 있음
 - TaP은 예측 결과가 오탐 없이 이상징후를 찾아내는 지를 나타내고, TaR은 얼마나 다양한 공격 범위를 찾아내는지를 나타냄

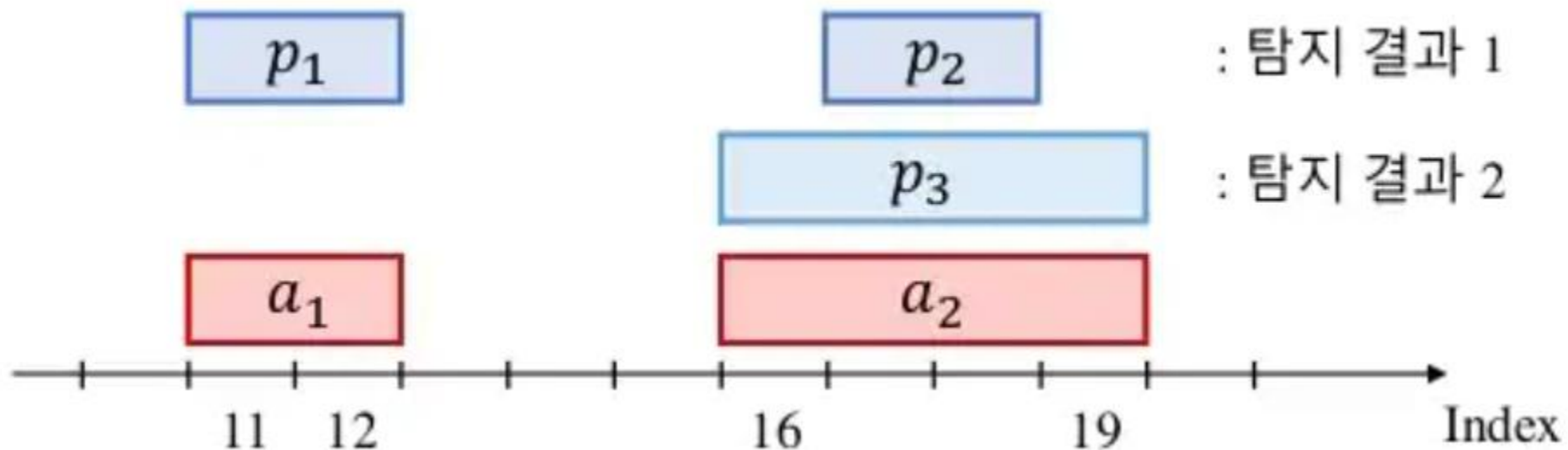


예측 결과	TaP	TaR
p_1	High	High
p_2	High	Low
p_3	Low	High
p_4	Low	Low

평가 지표

- TaPR의 평가 목표

- **목표 1. 탐지된 공격의 다양성 평가**
- 목표 2. 탐지의 정확성
- 목표 3. 낮은 오탐

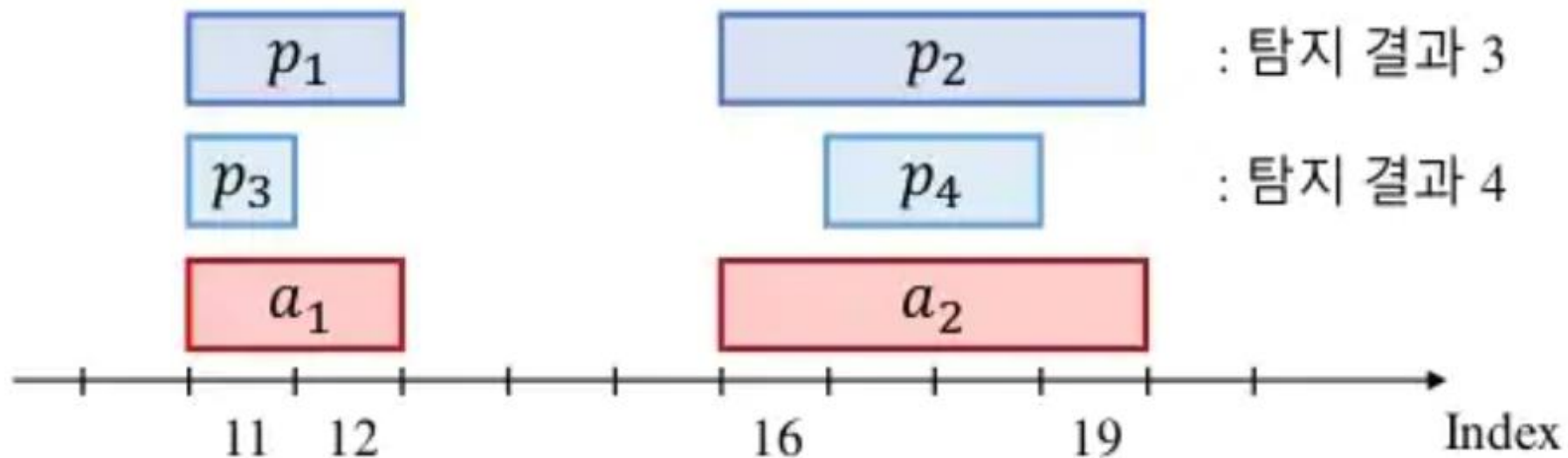


결과 1이 더 높게 평가 됨

평가 지표

■ TaPR의 평가 목표

- 목표 1. 탐지된 공격의 다양성 평가
- **목표 2. 탐지의 정확성**
- 목표 3. 낮은 오탐

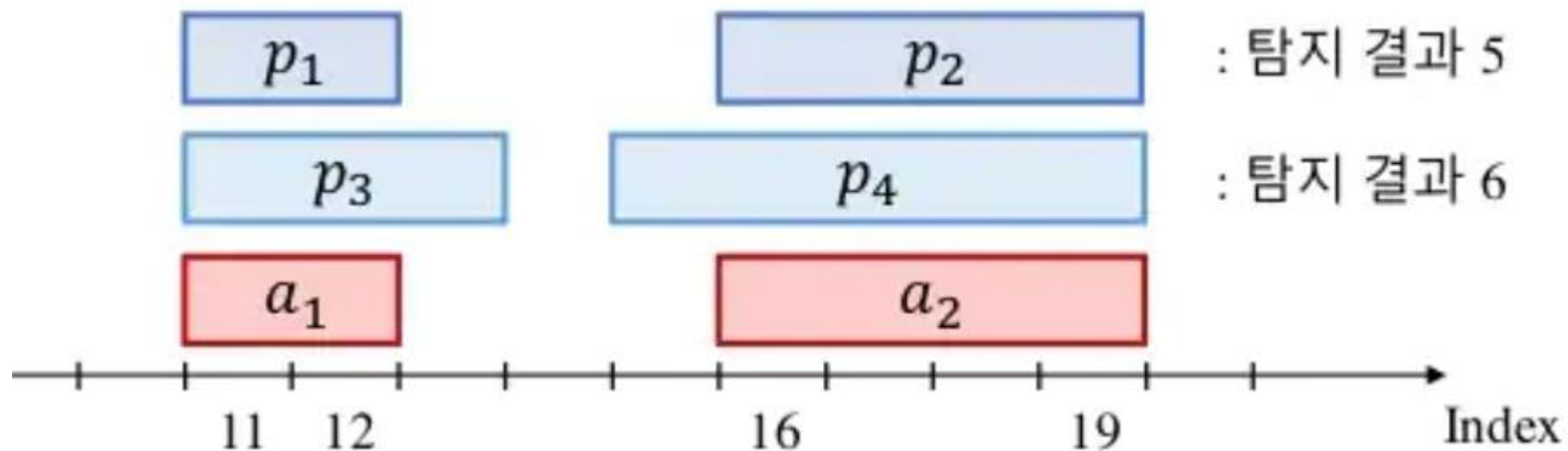


결과 3이 더 높게 평가 됨

평가 지표

■ TaPR의 평가 목표

- 목표 1. 탐지된 공격의 다양성 평가
- 목표 2. 탐지의 정확성
- **목표 3. 낮은 오탐**



결과 5가 더 높게 평가 됨