

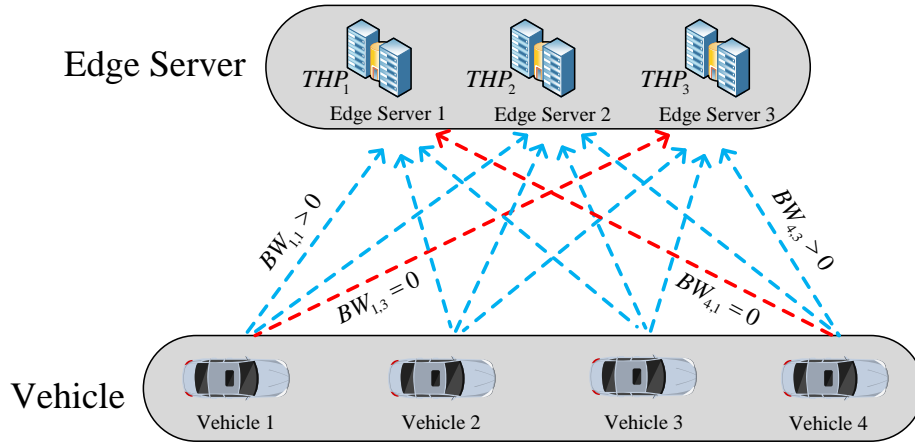
# Appendix

Anonymous Authors

Anonymous Institution

## Appendix-1

### Semantic Information Shunt Transmission Mode



**Fig. 1.** Semantic information shunt transmission mode.

In the semantic communication-based IoV crowdsensing tasks, we consider a more flexible and efficient data shunt transmission model, as shown in Fig.1, which allows users to offload data to multiple edge servers with different bandwidths. This shunt transmission mode can greatly increase the data transmission rate.

### The Training Algorithm for SVRANet

Algorithm 1 represents the training algorithm of SVRANet, where line 1 initializes the parameters (the learning rate, Lagrange multipliers, Lagrange multiplier period, etc.),  $\alpha$  denotes the learning rate for the gradient ascent process that maximizes the user misreporting, and  $\beta$  denotes the learning rate for the gradient descent process that minimizes the loss function. When training SVRANet for  $T$  epochs, each epoch uses a minibatch dataset  $S_t \subset S$ , and 4-7 lines are subjected to gradient ascent to find the users' optimal misreporting  $rgt_i^{(l)}(w^t)$ . Then, lines 8-9 execute gradient descent to reduce the loss function and update the network parameters. Finally, the Lagrange multipliers are updated.

---

**Algorithm 1** The training algorithm for SVRANet

---

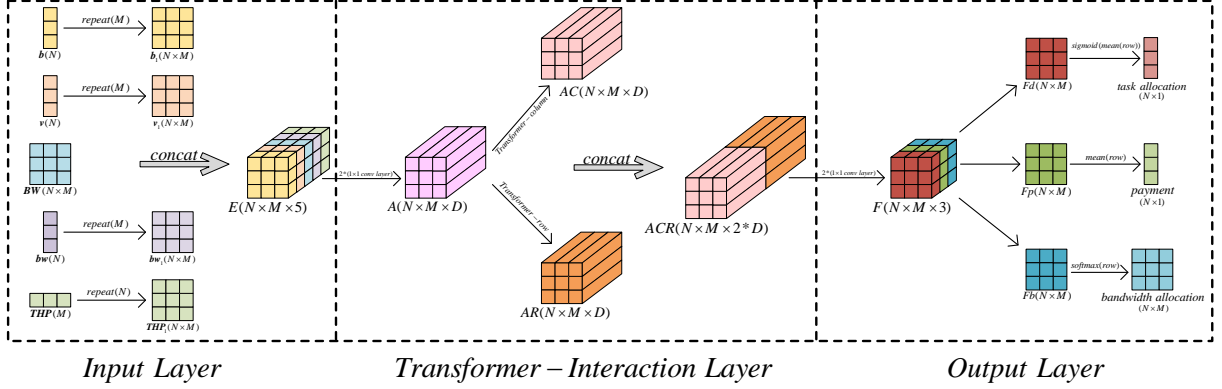
**Input:** Minibatches  $S_1, S_2, \dots, S_T$ ,  $S_t \subset S$ ,  $t \in [0, T]$ , Budget  $B$

- 1: **Initialize:**  $\forall \rho_\lambda > 0, \rho_\mu > 0, \rho_\eta > 0, \rho_\delta > 0, \rho_\gamma > 0,$   
 $T \in \mathbb{N}, R \in \mathbb{N}, T_\lambda \in \mathbb{N}, T_\mu \in \mathbb{N}, T_\eta \in \mathbb{N}, T_\delta \in \mathbb{N}, T_\gamma \in \mathbb{N}$   
 $w^0 \in \mathbb{R}^d, \lambda^0 \in \mathbb{R}^N, \mu^0 \in \mathbb{R}^N, \eta^0 \in \mathbb{R}, \delta^0 \in \mathbb{R}^{N \times M}, \gamma^0 \in \mathbb{R}^M.$
- 2: **for**  $t = 0$  to  $T$  **do**
- 3:   Receive minibatch  $S_t$
- 4:   Initialize false reports  $\mathbf{b}_i^{(l)} \in F_i, \forall l \in S_t, i \in N$
- 5:   **for**  $r = 0$  to  $R$  **do**
- 6:      $\forall l \in S_t, i \in \mathcal{N} : \mathbf{b}_i^{(l)} \leftarrow \mathbf{b}_i^{(l)} + \alpha \nabla_{\mathbf{b}_i^{(l)}} u_i^w((\mathbf{b}_{-i}, b_i))$
- 7:   **end for**
- 8:   Compute  $rgt(w)$ ,  $irp(w)$ ,  $bfp(w)$ ,  $bwp(w)$ , the gradient  $thpp(w)$  and the gradient update  $w^t$ .
- 9:    $w^{t+1} \leftarrow w^t - \beta \nabla_w \mathcal{L}_{\rho_\lambda, \rho_\mu, \rho_\eta, \rho_\delta, \rho_\gamma}(w^t, \lambda^t, \mu^t, \eta^t, \delta^t, \gamma^t)$
- 10:   Update the Lagrange multipliers once in several iterations:
- 11:   **if**  $t$  is a multiple of  $T_\lambda$  **then**
- 12:      $\lambda_i^{t+1} \leftarrow \lambda_i^t + \rho_\lambda rgt_i(w^{t+1}), \forall i \in \mathcal{N}$
- 13:   **else**
- 14:      $\lambda_i^{t+1} \leftarrow \lambda_i^t$
- 15:   **end if**
- 16:   **if**  $t$  is a multiple of  $T_\mu$  **then**
- 17:      $\mu_i^{t+1} \leftarrow \mu_i^t + \rho_\mu irp_i(w^{t+1}), \forall i \in \mathcal{N}$
- 18:   **else**
- 19:      $\mu_i^{t+1} \leftarrow \mu_i^t$
- 20:   **end if**
- 21:   **if**  $t$  is a multiple of  $T_\eta$  **then**
- 22:      $\eta^{t+1} \leftarrow \eta^t + \rho_\eta bfp(w^{t+1})$
- 23:   **else**
- 24:      $\eta^{t+1} \leftarrow \eta^t$
- 25:   **end if**
- 26:   **if**  $t$  is a multiple of  $T_\delta$  **then**
- 27:      $\delta_{i,j}^{t+1} \leftarrow \delta_{i,j}^t + \rho_\delta bwp_{i,j}(w^{t+1}), \forall i \in \mathcal{N}, j \in \mathcal{M}$
- 28:   **else**
- 29:      $\delta_{i,j}^{t+1} \leftarrow \delta_{i,j}^t$
- 30:   **end if**
- 31:   **if**  $t$  is a multiple of  $T_\gamma$  **then**
- 32:      $\gamma_j^{t+1} \leftarrow \gamma_j^t + \rho_\gamma thpp_j(w^{t+1}), \forall j \in \mathcal{M}$
- 33:   **else**
- 34:      $\gamma_j^{t+1} \leftarrow \gamma_j^t$
- 35:   **end if**
- 36: **end for**

---

## Appendix-2

### The Architecture of SVRANet



**Fig. 2.** The architecture of SVRANet.

SVRANet consists of 3 parts: *Input Layer*, *Transformer – Interaction Layer* and *Output Layer*. The detailed descriptions of these 3 parts are as follows.

**Input Layer** The inputs of the Input Layer are bids  $\mathbf{b} \in \mathcal{R}_{\geq 0}^N$ , values  $\mathbf{v} \in \mathcal{R}_{\geq 0}^N$ , bandwidth requirements  $\mathbf{bw} \in \mathcal{R}_{\geq 0}^N$ , channel capacities  $\mathbf{BW} \in \mathcal{R}_{\geq 0}^{N \times M}$ , and edge server throughput levels  $\mathbf{THP} \in \mathcal{R}_{\geq 0}^M$ . We extend the sizes of  $\mathbf{b}$ ,  $\mathbf{v}$ ,  $\mathbf{bw}$  and  $\mathbf{THP}$  to  $N \times M$ , obtaining  $\mathbf{b}_1 \in \mathcal{R}_{\geq 0}^{N \times M}$ ,  $\mathbf{v}_1 \in \mathcal{R}_{\geq 0}^{N \times M}$ ,  $\mathbf{bw}_1 \in \mathcal{R}_{\geq 0}^{N \times M}$  and  $\mathbf{THP}_1 \in \mathcal{R}_{\geq 0}^{N \times M}$ :

$$\begin{aligned} \mathbf{b}_1 &= \text{repeat} - \text{column}(\mathbf{b}, M) \in \mathcal{R}_{\geq 0}^{N \times M}, \quad \mathbf{v}_1 = \text{repeat} - \text{column}(\mathbf{v}, M) \in \mathcal{R}_{\geq 0}^{N \times M} \\ \mathbf{bw}_1 &= \text{repeat} - \text{column}(\mathbf{bw}, M) \in \mathcal{R}_{\geq 0}^{N \times M}, \quad \mathbf{THP}_1 = \text{repeat} - \text{row}(\mathbf{THP}, N) \in \mathcal{R}_{\geq 0}^{N \times M} \end{aligned}$$

Next, we concatenate  $\mathbf{b}_1$ ,  $\mathbf{v}_1$ ,  $\mathbf{bw}_1$ ,  $\mathbf{BW}$  and  $\mathbf{THP}_1$  to obtain a 3-dimensional tensor that contains global input information; this tensor is denoted as  $E \in \mathcal{R}^{N \times M \times 5}$  and is used as the input of the Transformer-Interaction Layer:

$$E = \text{concat}[\mathbf{b}_1; \mathbf{v}_1; \mathbf{bw}_1; \mathbf{BW}; \mathbf{THP}_1]$$

**Transformer-Interaction Layer** The input of the Transformer-Interaction Layer is a 3-dimensional tensor  $E \in \mathcal{R}^{N \times M \times 5}$  that contains global input information. We use two  $1 \times 1$  convolutions with a rectified linear unit (*ReLU*) activation function to increase the third dimension of  $E$  from 5 to  $D$  (the  $1 \times 1$  convolutions do not destroy the permutation-equivariant property), which is denoted as  $A \in \mathcal{R}^{N \times M \times D}$ ,  $\text{ReLU}(x) := \max(0, x)$ :

$$A = \text{conv}_2(\text{conv}_1(E)) \in \mathcal{R}^{N \times M \times D}$$

To obtain the task allocation scheme, payment scheme, and bandwidth allocation scheme, we let the edge server interact with its connected users via the Transformer to obtain  $AC_j \in \mathcal{R}^{N \times D}$ ,  $\forall j \in \mathcal{M}$ :

$$\begin{aligned} AC_j &= \text{Transformer} - \text{column}(A_{*,j}) \in \mathcal{R}^{N \times D}, \quad \forall j \in \mathcal{M} \\ AC &= \text{concat}[AC_1, AC_2, \dots, AC_M] \in \mathcal{R}^{N \times M \times D} \end{aligned}$$

We let the user interact with its connected edge servers via the Transformer to obtain  $AC_i \in \mathcal{R}^{M \times D}$ ,  $\forall i \in \mathcal{N}$ :

$$\begin{aligned} AR_i &= \text{Transformer} - \text{row}(A_{i,*}) \in \mathcal{R}^{M \times D}, \forall i \in \mathcal{N} \\ AR &= \text{concat}[AR_1, AR_2, \dots, AR_N] \in \mathcal{R}^{N \times M \times D} \end{aligned}$$

Finally, we concatenate the interaction result  $AC$  with  $AR$  to obtain a 3-dimensional tensor  $ACR \in \mathcal{R}^{N \times M \times 2 \times D}$  that contains the total mutual interaction information between the users and edge servers:

$$ACR = \text{concat}[AC; AR] \in \mathcal{R}^{N \times M \times 2 \times D}$$

The above material describes the network architecture of a single Transformer-Interaction Layer. We can stack multiple Transformer-Interaction Layers to enable SVRANet to acquire sufficient interaction information. When stacking multiple transformer interaction layers, we utilize two  $1 \times 1$  convolutions to reduce the third dimension of  $ACR$ . In the final layer, we reduce the third dimension to 3, obtaining a global feature map  $F \in \mathcal{R}^{N \times M \times 3}$  as the input for the Output Layer. Otherwise, the third dimension is reduced to  $D$  and is input into the Transformer-Interaction Layer for interaction purposes.

**Output Layer** The input of the Output Layer is the global feature map  $F \in \mathcal{R}^{N \times M \times 3}$ , and the task allocation scheme, payment scheme and bandwidth allocation scheme are calculated through  $F$ . As shown in the Output Layer in Fig. 2, we divide  $F$  into 3 parts,  $Ft \in \mathcal{R}^{N \times M}$ ,  $Fp \in \mathcal{R}^{N \times M}$  and  $Fd \in \mathcal{R}^{N \times M}$ , which are used to calculate the task allocation scheme, payment scheme and bandwidth allocation, respectively.

We calculate the mean of  $Ft \in \mathcal{R}^{N \times M}$  in a rowwise manner and then apply the *Sigmoid* activation function to ensure that task allocation is within  $(0, 1)$ ,  $\text{Sigmoid}(x) := \frac{1}{1+e^{-x}} \in (0, 1)$ :

$$a_i^w = \text{Sigmoid}(\text{mean}(Ft_{i,*}, \text{row})) \in (0, 1), \forall i \in \mathcal{N}$$

We calculate the mean of  $Fp \in \mathcal{R}^{N \times M}$  in a rowwise fashion to obtain the payment:

$$p_i^w = \text{mean}(Fp_{i,*}, \text{row}) \in (0, 1), \forall i \in \mathcal{N}$$

We apply the *Softmax* activation function in a rowwise manner to ensure that the bandwidth allocation level satisfies the user's requirements, with  $\text{Softmax}(x_j) := \frac{e^{x_j}}{\sum_{j=1}^M x_j} \in (0, 1)$ ,  $\sum_{j=1}^M x_j = 1$ :

$$f_{i,j}^w = \text{Softmax}(Fd_{i,j}, \text{row}) \in (0, 1), \forall i \in \mathcal{N}, j \in \mathcal{M}$$

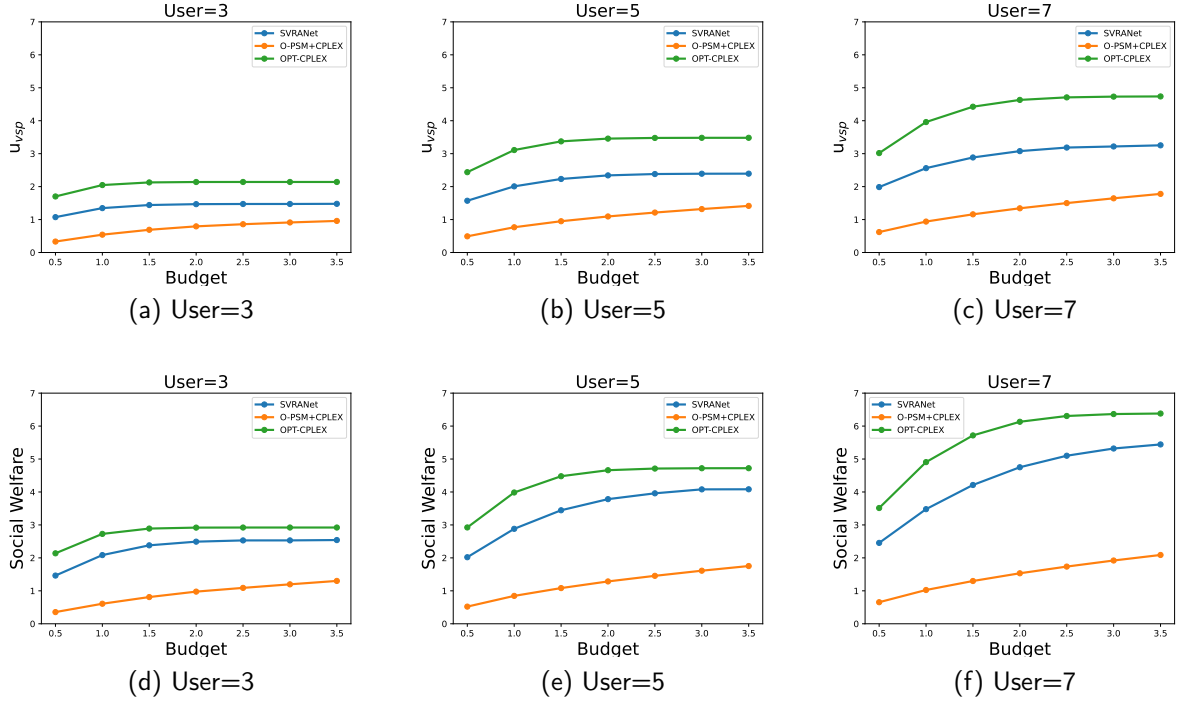
## Appendix-3

### The Impact of the Budget on Utility and Social Welfare at Different User Scales.

In terms of utility acquisition, SVRANet outperforms OPSM+CPLEX and can achieve over 65% of OPT-CPLEX, as shown in Fig. 3a,3b,3c. In terms of social welfare acquisition, SVRANet performs far better than OPSM+CPLEX and can achieve over 87% of OPT-CPLEX, as shown in Fig. 3d,3e,3f.

### Experimental data

As shown in Tables 1, 2, 3, 4, 5 and 6, we present the experimental data of  $u_{vsp}(w)$ ,  $rgt(w)$ ,  $irp(w)$ ,  $bfp(w)$ ,  $bwp(w)$  and  $thpp(w)$ .



**Fig. 3.** The impact of the budget on Utility and Social Welfare at different user scales.

**Table 1.**  $u_{vsp}(w)$

$user \backslash budget$	0.5	1.0	1.5	2.0	2.5	3.0	3.5
2	0.7706	0.9469	0.9836	0.9944	0.9936	0.9907	0.9929
3	1.0737	1.3474	1.4412	1.4678	1.4727	1.4732	1.4777
4	1.3417	1.6866	1.8526	1.9039	1.9257	1.9347	1.9355
5	1.5711	2.0077	2.2318	2.3407	2.3818	2.3912	2.3933
6	1.7802	2.3091	2.5581	2.7325	2.7979	2.8265	2.8261
7	1.9861	2.5604	2.8847	3.0773	3.1854	3.2185	3.2526

**Table 2.**  $rgt(w)$

$user \backslash budget$	0.5	1.0	1.5	2.0	2.5	3.0	3.5
2	0.000332	0.000278	0.000256	0.000216	0.000214	0.000222	0.000212
3	0.000331	0.000291	0.000260	0.000226	0.000211	0.000218	0.000218
4	0.000321	0.000309	0.000265	0.000241	0.000227	0.000211	0.000224
5	0.000307	0.000305	0.000276	0.000252	0.000227	0.000226	0.000225
6	0.000302	0.000295	0.000287	0.000268	0.000244	0.000229	0.000216
7	0.000297	0.000296	0.000295	0.000271	0.000262	0.000245	0.000229

**Table 3.**  $irp(w)$ 

$user \backslash budget$	0.5	1.0	1.5	2.0	2.5	3.0	3.5
2	0.000131	0.000191	0.000195	0.000183	0.000190	0.000179	0.000195
3	0.000120	0.000161	0.000192	0.000198	0.000182	0.000192	0.000187
4	0.000104	0.000160	0.000177	0.000187	0.000196	0.000196	0.000187
5	0.000098	0.000139	0.000167	0.000171	0.000185	0.000191	0.000186
6	0.000093	0.000129	0.000164	0.000174	0.000176	0.000183	0.000175
7	0.000083	0.000124	0.000147	0.000160	0.000179	0.000182	0.000187

**Table 4.**  $bfp(w)$ 

$user \backslash budget$	0.5	1.0	1.5	2.0	2.5	3.0	3.5
2	0.000106	0.000082	0.000065	0.0	0.0	0.0	0.0
3	0.000140	0.000097	0.000085	0.000063	0.000046	0.0	0.0
4	0.000144	0.000128	0.000100	0.000091	0.000072	0.000061	0.000035
5	0.000156	0.000130	0.000113	0.000112	0.000090	0.000075	0.000064
6	0.000163	0.000160	0.000135	0.000124	0.000113	0.000093	0.000060
7	0.000184	0.000178	0.000169	0.000140	0.000135	0.000119	0.000072

**Table 5.**  $bwp(w)$ 

$user \backslash budget$	0.5	1.0	1.5	2.0	2.5	3.0	3.5
2	0.0000562	0.0000583	0.0000577	0.0000522	0.0000588	0.0000576	0.0000654
3	0.0000514	0.0000581	0.0000585	0.0000596	0.0000654	0.0000610	0.0000656
4	0.0000505	0.0000579	0.0000535	0.0000594	0.0000639	0.0000605	0.0000591
5	0.0000471	0.0000550	0.0000578	0.0000609	0.0000572	0.0000573	0.0000599
6	0.0000462	0.0000519	0.0000553	0.0000570	0.0000574	0.0000548	0.0000557
7	0.0000428	0.0000505	0.0000516	0.0000572	0.0000607	0.0000566	0.0000542

**Table 6.**  $thpp(w)$ 

$user \backslash budget$	0.5	1.0	1.5	2.0	2.5	3.0	3.5
2	0.000035	0.000044	0.000047	0.000043	0.000049	0.000048	0.000046
3	0.000050	0.000060	0.000063	0.000063	0.000066	0.000072	0.000070
4	0.000056	0.000068	0.000070	0.000080	0.000082	0.000081	0.000082
5	0.000060	0.000078	0.000081	0.000082	0.000083	0.000093	0.000095
6	0.000064	0.000074	0.000086	0.000092	0.000096	0.000102	0.000108
7	0.000068	0.000083	0.000093	0.000113	0.000116	0.000119	0.000115