Workgroup:Network Working Group
Internet-Draft:
draft-zhang-rvp-problem-statement-00
Published:20 October 2025
Intended status: Informational
Expires:20 April 2026

X.Zhang
Hicagent
Y.Zhang
C.Zhang
M.Zhang
China Telecom
W.Lu
SUPCON

Problem Statements and Requirements of Real-Virtual Agent Protocol (RVP): Communication Protocol for Embodied Intelligence in Physical-Digital Continuum

Status of this Memo

This Internet-Draft will expire on April 20, 2026.

Abstract

The Real-Virtual Agent Protocol (RVP) enables seamless coordination between physical entities (robots, IoT devices, manufacturing systems and agents) and digital agents (AI systems, software agents, virtual twins) through unified composite identity management, physical/social/production relations graph-based coordination, and physical constraint integration. Unlike existing protocols that assume peer-to-peer digital relationships (A2A for agents, MCP for AI tools, ANP for agent networks), RVP unifies physical and digital agents communication and achieves physical data loop for online learning for embodied agents considering both hierarchical physical/social/production relations and physical world constraints. RVP is designed for immediate deployment in modern manufacturing, smart cities, autonomous mobility systems, and human-AI collaborative environments where non-peer, partially centralized relations and coordination is essential for real-world embodied intelligence networks.

Table of Contents

1. Introduction

1.1. Background and Current Status of Existing Protocols

The convergence of artificial intelligence, robotics, and ubiquitous computing is creating unprecedented demand for connectivity and coordination between physical systems and digital intelligence. However, existing protocols are insufficient to effectively meet these demands, facing severe challenges in cross-entity collaboration, cross-level management, and cross-architecture communication, primarily manifested in problems such as identity fragmentation, protocol incompatibility, and missing physical constraint reflection mechanisms.

With the rapid development of Large Language Models (LLMs), LLM-based agents have become very popular in many industries and scenarios. Agent and model protocols are gaining widespread attention from both industry and academia. Meanwhile physical agents(e.g., emobodied robots, autonomous driving cars, UAVs) embedded in physical entities are playing a crucial role in modern manufacturing[SMART-FACTORY], smart cities[AI-CITY], autonomous mobility systems[APPOLO], and human-AI collaborative environments[Human-Robot]. The interaction between hetergenous physcial agents and digital agents as well as models bridges the physical world and digital world into a seamless continuum. Real-Virtual Agent Protocol (RVP) is a communication protocol designed for reality-virtuality symbiotic, supporting non-peer-to-peer, and partially centralized embodied intelligence agent networks. It focuses on mapping on physical/social/production relations graphs and physical constraint reflection between agents.

Today’s agent communication protocols are designed for homogeneous environments. Mainstream agent communication protocols include A2A[A2A-SPEC], MCP[MCP-SPEC], ANP[ANP-SPEC], etc.:

Agent-to-Agent (A2A) Communication Protocol: Supports digital agent communication but assumes all participants are software entities operating in digital environments in peer-to-peer communication, hardly to handle non-peer relationships in physical world, and lacks of physical/social/production relations constraints. In practical applications, it has fundamental limitations for physical-digital coordination, including:

1) Communication Model Mismatch: A2A protocol uses asynchronous message transmission optimized for digital agents that can pause, queue, and process messages at variable speeds. Physical systems require real-time coordination with hard deadlines-a robot cannot "pause" mid-motion to wait for queued messages.

2) No Physical Constraint Modeling: A2A protocol has no standard way to represent physical constraints (workspace boundaries, inertia, safety requirements). When a digital agent requests a robot to "move to position X immediately," the protocol cannot express that the robot needs 3.2 seconds due to physical acceleration limits.

3) Safety Integration Gap: A2A protocol has no built-in safety mechanisms. There is no standard way to represent that certain agent communications could cause physical harm if mishandled.

MCP: Provides AI systems with tool access but treats physical systems as external, and sometimes stateless tools rather than intelligent twin participants. It focuses on model context management, concentrating on context interaction between agents and tools (such as user preferences, environmental states, session and task information, etc.). The problems raised by physcial systems include:

1)Stateless Tool Model: MCP treats all tools as stateless functions. Physical systems are inherently stateful, i.e., a robot’s current position affects what next operations are possible. MCP is hard to represent that the same tool call may succeed or fail based on physical context.

2)Bidirectional Coordination: MCP is bascially request-response model. Physical systems need to provide continuous feedback (sensor data, status updates, emergency conditions) that should influence AI decision-making in real-time.

3) MCP basically assumes one single AI system/agent accessing tools, which does not involve complex collaboration relationships between agents. Physical agents often need to coordinate between multiple AI systems (planning AI, safety monitoring AI, quality control AI) simultaneously. Moreover, real-time interaction and synchronization (e.g., ms magnitude) between physical and digital agents is hard to achieve.

Agent Network Protocol (ANP): Organizes agent networks in multi-agent scenarios with peer-to-peer entities, building the underlying architecture of the agent internet, known as the "HTTP of P2P networks" and supporting decentralized decision-making. It supports decentralized discovery but identity management is based on DID, unable to handle 1-to-N topology relations from the physical entity to virtual entities (for examaple, an automonous driving car involves a physcial entitiy and several virtual assistant entities to do route optimization and simulation, operate the on-car electronic systems and search for the weather forecast,etc.)and virtual entity cluster discovery requires multiple queries with low efficiency. Its constraints are similar to the A2A communication protocol and will not be elaborated further.

Additionally, regarding connecting physical and digital entities, existing underlying transport protocols have the following issues:

IoT Protocols (MQTT, CoAP[RFC7252], HTTP) handle device connectivity and data exchange but lack semantic understanding of physical-digital integration requirements.

IoT protocols like MQTT and CoAP handle device connectivity but cannot support intelligent coordination:

1) No Coordination Semantics: IoT protocols transport data but have no understanding of coordination requirements. Publishing sensor data to a topic provides no information about what coordination actions are appropriate.

2) No Identity Coherence: IoT protocols identify devices by topics or endpoints but cannot represent that multiple endpoints refer to the same logical entity. A robot’s position sensor, control interface, and status feed are separate MQTT topics with no protocol-level connection.

3) No Intelligence Integration: IoT protocols have no standard way to integrate with AI systems beyond basic pub/sub. There is no protocol support for AI systems to understand device capabilities or coordinate intelligent actions.

RVP is implemented similarly to existing protocols (such as MCP, A2A, ANP) through further encapsulation or modification of underlying transport protocols (such as SSE, WebSocket, MQTT) or even direct modification of A2A, MCP and ANP.

1.2. Why RVP is Necessary

Existing protocols (A2A, MCP, IoT protocols) have fundamental limitations in handling unified coordination of physical entities and digital entities. Although A2A protocols could theoretically be extended to handle physical entities, this would require fundamental changes to agent communication semantics to handle physical constraints [A2A-FIPA][CPS-DESIGN]and complete redesign of unified identity management across domains. Enhancing MCP to effectively handle physical and digital agent coordination would require new design principles to design stateful coordination and real-time communication among different kinds of agents, which is common in real world. It’s hard to imagine to convert all physical systems into stateless tools, thereby losing their needs of intelligent coordination. Adding intelligence to IoT protocols would require fundamental semantic changes incompatible with existing IoT deployments, still lacking unified identity management and limited to device-level rather than system-level coordination.

Future multi-agent architectures will evolve into hierarchical (MCP’s C/S architecture can satisfy), distributed (ANP’s P2P architecture can satisfy), and hybrid frameworks, enabling flexible and diverse ways to coordinate large numbers of agents to complete various complex tasks under different business scenarios. However, the existing architectures cannot fully meet the communication needs of reality-virtuality integrated embodied intelligence agent networks. For example, one-to-many real-virtual mapping requires simultaneous identity authentication of multiple virtual entities and simultaneous registration and discovery of these agents. Communication between real-virtual agents also requires following constraints and rules of production relations graphs in the physical world. What’s more, the data starvation for model training requires real-time physical data input to achieve data-loop for online training in embodied intelligence.

Therefore, RVP is needed that can treat physical and digital as expressions of unified entities and create efficient communication based on these expressions. Instead of replacing existing protocols, the new protocol is designed to integrate with them:

1) Integration with Existing Protocols: Through adapters or middleware layers, enabling the new protocol to work collaboratively with existing protocols such as A2A, MCP, IoT, etc.

2) Introducing New Mechanisms: Particularly physical constraint reflection, unified identity management, physical/social/production relations graphs, etc. Based on this, a heterogeneous communication protocol for reality-virtuality integrated embodied intelligence agents is proposed.

2. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

Composite Agent: A Composite Agent is a super-agent to interact with the external physical world on behalf of its associated Real-entity Agent, and Virtual-entity Agent. While adhering to the principle of permission isolation, it realizes effective utilization and comprehensive fusion of information related to both worlds. By abstractly layering the capabilities of individual entities in the physical world, it achieves efficient coordination from value decision-making to task execution. Composite Agent aims to characterize human/unmaned behavior patterns in complex environments, ensuring efficiency, flexibility, and stability of task processing.

Physical Expression: The concrete instance of an object in the physical domain, carrying the existence form and interaction capabilities of the object in the physical world.

Digital Expression: The virtualized instance of an object in the digital domain, carrying the logical representation and computational capabilities of the object in digital space.

Real-Virtual Agent Protocol (RVP): A standardized communication protocol specifically designed for composite agent architecture, regulating interactions and coordination among physical and digital agents.

Physical/Social/Production Relations Graph: A semantic network-based directed graph describing organizational relationships, business processes, and permission constraints among agents.

Physical Constraint: Standardized description of timing, spatial, dynamic, and safety constraints in the physical world, supporting automatic verification and real-time reflection.

3. Problem Statement

3.1. Identity Fragmentation Across Domains

Currently, in above mainstream communication protocols, no protocol provides unified identity management across physical and digital domains. Each protocol treats the same entity as a separate, unrelated object. The same logical entity exists as disconnected representations in different systems. These representations are manually coordinated through custom integration code, leading to consistency problems and coordination failures.

Without unified identity management, the same entity has multiple un-associated identity representations in different systems, leading to low registration efficiency, data inconsistency, and coordination failures. Identity information needs to be manually synchronized, increasing system complexity and maintenance costs.

3.2. Lack of Physical/Social Relations Mapping among agents

Current mainstream agent protocols (MCP/A2A/ANP) are built on a digital-native worldview, ignoring the inherent non-peer nature, hierarchical structures, and business constraints of agents in the physical world. In existing agent systems, interactions between agents are often based on simple peer-to-peer or client-server models, lacking mapping of complex production relationships in the real world.

In real world, an agent may represent the individual person’s social roles in different scenarios, such as father, employee, friend, etc.; it can also represent actual existence in the physical world, such as robots, sensors, actuators, and other hardware devices. Collaboration in the real world is often non-peer, with hierarchical structures, division of labor, and permission constraints. For example, in intelligent manufacturing, there are levels such as scheduling centers, production line control, and robots; in smart cities, there are command centers, departments, and field devices.

Therefore unlike in digital world, the hierarchical and network-based physical/social/production relations graph exists, which stores physical world data and capabilities related to specific roles.

Due to the lack of explicit modeling and mapping of these relationships, it leads to:

Difficult Permission Control: Access and operation permissions between agents are not constrained based on physical/social/production relations, potentially causing unauthorized operations.

Low Collaboration Efficiency: Agents cannot quickly identify collaboration partners, requiring complex negotiation and discovery mechanisms.

Unreasonable Resource Allocation: Due to the lack of relations-based scheduling, resource allocation may not conform to actual business logic, leading to resource waste or bottlenecks.

Poor System Maintainability: When business relations change, manual adjustments to connections and permissions between agents are needed, which is error-prone.

Therefore, we need to introduce a "Phycial/Social/Production Relations Graph" in the protocol to explicitly describe these relations among agents and implement permission control, collaboration scheduling, and resource allocation based on this graph.

3.3. Lack of Physical Constraint Reflection

The current agent protocol ecosystem has a fundamental cognitive disconnect: intelligent protocols like A2A and MCP can understand intelligent coordination but lack physical perception. In a research [STI-BENCH] it pointed out While LLMs have been extensively studied for visual semantic understanding tasks, their ability to perform precise and quantitative spatial-temporal understanding in real-world applications is rather week, with top-performing models like Gemini-2.5-Pro achieving around 41.4% average accuracy.

IoT protocols can understand physical devices but lack intelligent cognition. This causes agent decisions in the digital world to be seriously disconnected from physical world reality rules.

Physical and digital spaces need to establish a standardized description framework for physical constraints to implement real-time reflection mechanisms for constraint states, ensuring consistency between digital decisions and physical reality.

3.4. Lack of Efficient Registration and Effective Discovery Mechanisms for Agents

Existing protocols (such as MCP, A2A, ANP) have deficiencies in registration and discovery, resulting in agents requiring manual configuration and inability to dynamically adapt to network changes. Due to the lack of unified, adaptive, semantic registration and discovery mechanisms, agent networks become information silos, unable to achieve true "plug-and-play" and "dynamic collaboration." A unified registration and discovery mechanism needs to be designed to support dynamic joining and exiting of agents and automatic updating of network topology.

3.5. Lack of Real-time Data Input to the Embodied Large Model for Training

Back to 2023, just after one year of ChatGPT’s release, many studies warned that the world could run out of high-quality data to train AI soon. The research[DATA-RUNOUT] predicted that the training dataset will run out between 2026 and 2032, or even earlier.

This is even worse for embodied intelligence systems, which has less recored knowledge to learn. Online learning from real-time practice data attracts more and more attentions to solve this problem. Particularly, data automatically generated during agent task execution that is not directly explicitly recorded contains deep information about environmental dynamics, physical laws, task execution context, decision-making processes, and multi-modal interactions. For example, in sanitation scenarios, implicit data includes interactions between sanitation vehicles and the environment (such as the relationship between cleaning force and ground material), micro-adjustments during task execution (such as trajectory corrections when avoiding pedestrians), and correlations between equipment operating status and work effectiveness. These real-time data supplement is vital for training the embodied agents. Moreover it will help to solve the data-training-decision fragmentation at the same time, i.e., when training data is historical, static, and disconnected, while agents need to make immediate responses in dynamic, real-time, continuous environments. This fragmentation prevents models from continuously evolving in the real world.

To the best of our knowledge, current protocols are not available to input real-time data for continuous training.

To summary, there is an urgent need for unified communication protocol standards to effectively connect physical entities and digital agents.

4. Protocol Requirements Discussion and Design Principles

4.1. Design Principles

Constraint Priority Principle: Physical constraints are not afterthought limitations but prerequisites for system design and operation.

Perform physical feasibility verification before all decisions and executions

Establish automatic propagation and consistency maintenance of constraints in task chains

Deeply integrate safety constraints into coordination semantics

Unified Semantics Principle: Establish a unified semantic framework across all protocols and domains, eliminating ambiguity in concept mapping.

Ensure consistent interpretation of the same concepts across different protocols and domains

Semantic interpretation considering current physical environment and business context

Support runtime semantic adjustment and optimization

4.2. Requirements Discussion

This section is intended to stimulate the open discussion for the RVP requirements, so that the design team can consolidate the requirements of the RVP protocol.

RVP.REQ-1: Unified Entity Identity Management

Addressing the identity fragmentation problem in existing protocol ecosystems, where the same entity has different identity marking systems in physical, digital, and hybrid domains, leading to semantic inconsistency, life cycle disconnection, and accumulated security risks. The protocol MUST support unified entity registration, containing physical capabilities, digital intelligence, and hybrid manifestations within a single identity model.

The requirements include:

Support automatic association and verification of physical and digital identities

Establish a globally unique entity identifier system

Identity registration MUST contain complete composite identity description.

Identity verification SHOULD support cross-domain identity consistency checks.

Identity updates MUST ensure synchronization of identity information in all domains

RVP.REQ-2: Dynamic Registration and Discovery Mechanisms

Existing systems rely on manual configuration and hard-coded connections, unable to adapt to dynamic network topology changes, resulting in poor system scalability. The protocol MUST provide efficient agent registration and service discovery mechanisms.

The requirements include:

New nodes MUST be able to automatically discover the network and complete registration

Node capability changes MUST be updated to the registry in real-time

Provide real-time node health status monitoring

Support external interface for natural language semantic service discovery

Service discovery MUST consider current context and environmental state

Provide intelligent service selection based on multi-dimensional matching

RVP.REQ-3: Physical/Social/Production Relations Graph Management

Existing protocols lack of expressive capability for real-world organizational structures, unable to express business relations such as command, collaboration, and affiliation, resulting in low coordination efficiency and coarse-grained permission control. The protocol MUST provide standard mechanisms to define, maintain, and query organizational structures among entities based on production relationships.

The Graph Model Requirements include:

Define standardized physcial/social/production relations classifications

Support runtime creation, update, and deletion of relations

Implement automatic derivation of permissions based on relations graphs

The Operational Requirements include:

Node registration MUST automatically derive initial physical/social/production relations

Relations changes SHOULD trigger impact analysis and system adjustments

Routing decisions MUST consider physical/social/production relations constraints

RVP.REQ-4: Physical Constraint Integration

Existing protocols lack systematic description and verification mechanisms for timing, spatial, dynamic, and other constraints, causing digital decisions to be disconnected from physical reality. The protocol MUST provide standard mechanisms to represent and enforce physical constraints in coordination decisions.

The requirements include:

Define machine-readable physical constraint description format, forming a standardized constraint language

Establish complete constraint classification

Constraint description MUST support automatic verification

Decision verification MUST perform physical constraint checks before execution

Constraint violations SHOULD trigger standardized emergency responses

Constraint propagation MUST maintain consistency during task decomposition

RVP.REQ-5: Safety-Aware Coordination Semantics

Safety mechanisms considerations are required to be added as built-in features to meet real-time requirements. The protocol MUST integrate physical safety constraints, timing requirements, and emergency response mechanisms into coordination communications.

The requirements include:

Built-in safety semantics, with safety constraints as core components of coordination semantics

Safety decisions consider physical state and business context

Operation execution SHOULD perform real-time safety boundary checks

Emergency response MUST be automatically triggered when safety threats are detected

RVP.REQ-6: Real-time Data Sharing

Real-time data sharing involve injecting real-time data of the physical agents into models and sharing direct and local sense, observation or other kind of data between agents.

The requirements include:

Extremely Low latency reliable data transmission

Streamed training data provision to training procedure

Training and reasoning mode switch dynamically

Model update and versioned distribution

Standard gradient/parameter aggregation interface

Direct sense and information sharing between agents

4.3. Protocol Integration Strategy

RVP is designed as a meta-protocol that does not replace existing protocols (such as A2A, MCP, IoT protocols) but inherits, extends, or enhances them, providing core capabilities such as unified entity identity, physical/social/production relations graphs, and physical constraint integration. A2A is the most suitable starting-point when designing RVP.

A2A Protocol Enhancement: Introduce RVP’s core concepts into the A2A protocol to support physical-digital unified coordination.

Extend Agent Card: Add composite identity description, dynamic capability discovery, and physical/social/production relations graph information to A2A’s Agent Card to support intelligent routing.

Unified Identity Mapping: Extend agent identity in A2A to unified entity identity (including Real-entity agent and Virtual-entity agent), achieving cross-domain identity consistency.

Real-time Coordination Semantics: Define new message types and interaction patterns for time-sensitive operations, supporting deadlines and real-time guarantees.

Physical Constraint Representation: Add physical constraint fields to A2A messages, allowing agents to consider physical limitations (such as location, capabilities, time, etc.) during negotiation.

Safety-aware Communication: Integrate safety context into A2A communication, including permission verification based on production relationships and physical safety constraints.

Besides A2P, MCP and IoT protocols can also integrate with RVP, but this is not mandatory in the 1st phase:

MCP Integration Enhancement: Enable MCP to support stateful, real-time coordination of physical agents.

Hardware Tool Extension: Extend MCP tools from predominantly software digital domain to hardware physical domain, supporting embodied intelligence agent communication for hardware device queries, instructions, and interactions.

Stateful Tool Representation: Encapsulate physical agents as stateful MCP tools, allowing tools to maintain state (such as device status, task progress).

Timing Constraint Specification: Add timing constraints (such as start time, deadline, execution duration, etc.) in MCP tool calls.

Multi-AI Coordination Support: Through RVP’s physical/social/production relations graph, coordinate multiple AI model operations on the same physical system.

IoT Protocol Bridge: Seamlessly integrate IoT devices into the RVP ecosystem, achieving both semantics of device data and intelligent coordination of device capabilities.

Semantic Mapping: Map device data (such as sensor readings) to coordination context (such as events, states) and attach physical constraints.

Unified Entity Identity: Assign unified entity identity to each IoT device and associate it with the corresponding Real-entity agents.

Device Capability Representation: Describe device capabilities as standardized services, including functions, constraints, and states.

Message Delivery Guarantees: Provide reliability guarantees (such as at-least-once delivery, acknowledgment mechanisms) for coordination messages among IoT devices.

5. RVP Use Cases

This section is not a to-do list for the protocol. To make RVP well understood, it provides several scenarios on how RVP could be used in practice.

5.1. Flexible Production Line Coordination in Manufacturing

A flexible manufacturing line produces customized products with different kinds of agents among robots, CNC machines, AGVs (Automated Guided Vehicles), and quality inspection stations, along with their digital twins [DIGITAL-TWINS] agents.

A robot arm Composite Agent is composed of the controller, Real-entity robot arm Agent and its digital twins (Virtual-entity Agent). When the composite robot arm agent is registered via RVP, the associated real-entity robot and virtual-entity agents are automatically registered as well, reducing the complexity of registration several times.

The Physical/Social/Production Relations Graph with hierarchical structure is created and registered which may look like: Production Line Manager Work Cell Controller Individual Robots. Permission constraints ensure only authorized controllers can issue commands to specific robots.

Physical Constraints are transferred via RVP with spatial constraints such as robot workspace boundaries, collision avoidance zones; temporal constraints such as task completion deadlines and synchronization requirements; and dynamic constraints such as Maximum acceleration, payload capacity, energy consumption, etc....

When the order is received, the production line manager agent decomposes tasks based on Physical/Social/Production Relations Graph and message routing is based on the relations graph, supporting explicit routing paths and automatic path calculation. The tasks in associated agents are linked based on the relations graph constraints supporting synchronous and asynchronous communication. The real-entity agents execute the appointed tasks with real-time monitoring and the virtual-entity agents simulate execution feasibility with physical constraints. The physical verification ensures digital instructions comply with physical system limitations after the tasks are completed. Constraint violation triggers automatic redo. The feedback as well as the process data in different agents can be integrated to the owner for decision-making and model training. Problematic nodes request assistance via relations graph queries using RVP.

5.2. Disaster Rescue Human-Robot Collaborative Environments

Consider a search and rescue operation with human rescuers, rescue robots, drones, and command center.

Robots and humans form ad-hoc coordination network in infrastructure-degraded environment. Flexible social relations graph adapting to situation is created via RVP to establish command structure including incident commander, who provides overall coordination; team leaders who coordinate local resources and Individual robots/humans, who have autonomy for immediate safety decisions. The social relations graph will dynamically change to add emergency relations (e.g., fire truck and ambulance).

The robot Composite Agents register their capabilities via RVP and the search area is decomposed based on the robot capabilities, human responsibility and environmental constraints to create several search teams(including humans and robots). The physical constraint is monitored for structural integrity or hazardous conditions. The virtual-entity agents simulate evacuation routes with physical constraint verification. Once the constraint violation is detected, the digital twin agent changes its simulation and automatic evacuation action is executed via the social relations graph. Dynamic re-planning and re-verification is executed as new constraints are discovered (e.g., aftershock occurs or blocked route) for the real and virtual agents. Real-time data from operation feeds embodied disaster response model.

5.3. Autonomous Driving Systems in Smart City

5.3.1. Multi-Vehicles(Agents) Collaborative Delivery

The fleet of autonomous delivery robots and drones across urban environment with dynamic obstacle avoidance and resource sharing is a multi-agents coordination. Each delivery vehicle is a Composite Agent including: high-level mission planner, the physical vehicle with sensors/actuators and virtual-entity agent for route optimization and simulation. The production Relations graph is like: Fleet management center coordinating individual vehicles peer relations between vehicles for resource negotiation hierarchical override for emergency situations.

The virtual-entity agent simulates route with physical feasibility check. The production relations graph queries nearby vehicles for resource sharing opportunities and coordination sessions are established. The real-entity agent executes with continuous constraint monitoring and trigger dynamic re-planning when when the constraint violates (e.g., unexpected obstacle, battery drain, etc...). The execution data (actual vs. predicted energy consumption, obstacle patterns) feeds the autonomous driving model via RVP.

5.3.2. Vehicle-to-Everything (V2X) Intersection Management

Consider an intelligent intersection with autonomous V2X vehicles, human-driven V2X vehicles, cyclists, and pedestrians without traditional traffic signals.

Once the vehicle approaches intersection, it registers it vehicle capabilities (e.g., size, braking distance, speed range and V2X information) and discovers intersection controller via RVP. The intersection controller establishes coordination sessions based on the physical relations graph among the vehicles. The virtual-entity agents simulate safe crossing sequences with physical constraint verification where physical safety constraints are the primary protocol feature to be checked, not afterthought. The Safety constraints like minimum time gaps, line-of-sight requirements, pedestrian priority are checked. The V2X vehicle, cyclist, pedestrian with smartphone receive crossing permissions with timing constraints via RVP. Continuous safety monitoring is executed and constraint violations trigger RVP protocols usage.

6. Security Considerations

6.1. Identity Verification and Authorization

Composite Identity Verification Not only verify digital identity but also verify physical identity, ensuring authenticity of unified identity.

Digital Identity: OAuth 2.0 tokens, API keys

Physical Identity: Device fingerprinting, TPM attestation, secure elements

Composite Binding: Cryptographic binding between physical and digital credentials

Multi-factor Authentication for critical operations, require multi-factor authentication, including physical tokens or biometrics tokens:

Level 1 (routine operations): Single factor (token/certificate)

Level 2 (sensitive operations): Two factors (token + OTP)

Level 3 (critical operations): Three factors (certificate + biometric + physical token)

Authorization Based on Relations Graph

Utilize physical/social/production relations graph for dynamic permission calculation, considering relations type, path length, time validity, etc.

6.2. Communication Security

End-to-End Encryption

Use TLS 1.3 or higher for transport encryption, and apply application-layer end-to-end encryption for sensitive data.

Message Integrity

Use digital signatures to ensure message integrity and prevent tampering.

6.3. Safety System Protection

Independent Safety Channels

Physical safety systems (such as emergency stop) use communication channels independent of coordination protocols.

Physical-Digital Boundary Protection

Deploy safety gateways between physical and digital domains for protocol filtering and deep inspection.

Gateway functions include:

Message content inspection and validation

Rate limiting and anomaly detection

Whitelist-based operation filtering

Constraint verification before physical execution

6.4. Audit and Monitoring

Comprehensive Logging Record: Detailed logs of all coordination sessions, including message content, participants, timestamps, etc.

6.5. Isolation Capabilities

Dynamic Isolation: Based on security events, dynamically isolate affected entities or coordination domains.

7. IANA Considerations

This document requests IANA to establish new registries for RVP protocol parameters. More details are in further versions.

8. References

8.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

8.2. Informative References

[A2A-SPEC] https://a2a-protocol.org/dev/specification/

[A2A-FIPA] Foundation for Intelligent Physical Agents, "FIPA Agent Communication Language Specifications", 2002.

[MCP-SPEC] Anthropic, "Model Context Protocol Specification", 2024.

[ANP-SPEC] "Agent Network Protocol Technical White Paper", https://arxiv.org/pdf/2508.00007, 2025.

[RFC7252] Shelby, Z., Hartke, K., and C. Bormann, "The Constrained Application Protocol (CoAP)", RFC 7252, June 2014.

[SMART-FACTORY] https://www.globaltrademag.com/smart-factories-how-technology-is-revolutionizing-manufacturing/.

[AI-CITY] https://www.deloitte.com/content/dam/assets-shared/docs/industries/government-public-services/2025/ai-powered-cities-of-the-future.pdf.

[APOLLO] https://www.apollo.auto/en/apollo-self-driving.

[Human-Robot] Dong, W., "Toward Embodied Intelligence-Enabled Human-Robot Symbiotic Manufacturing: A Large Language Model-Based Perspective",Journal of Computing and Information Science in Engineering, May 2025.

[DIGITAL-TWINS] Grieves, M., "Digital Twin: Manufacturing Excellence through Virtual Factory Replication", Digital Manufacturing, 2014.

[CPS-DESIGN] Lee, E., "Cyber Physical Systems: Design Challenges", University of California, Berkeley Technical Report, 2008.

[STI-BENCH] Li, Y., "STI-Bench:Are MLLMs Ready for Precise Spatial-Temporal World Understanding?", ICCV 2025.

[DATA-RUNOUT] Villalobos,P.,Will we run out of data? Limits of LLM scaling based on human-generated data, https://arxiv.org/pdf/2211.04325.

9. Acknowledgments

The authors thank Dirk Dressler who provided valuable feedback during the development of this specification.

Authors' Addresses

Xiaoxun Zhang
Hicagent
No.57 Boxia Road, Pudong Software Park,Shanghai, P.R.China
zhangxiaoxun@hicagent.com

Yunfei Zhang(Editor)
China Telecom
No.31 Jinrong Street, Xicheng District,Beijing, P.R.China
Zhangyf80@chinatelecom.cn

Chi Zhang
China Telecom
No.31 Jinrong Street, Xicheng District,Beijing, P.R.China
Zhangc120@chinatelecom.cn

Min Zhang
China Telecom
No.31 Jinrong Street, Xicheng District,Beijing, P.R.China
Zhangmin3@chinatelecom.cn

Weijun Lu
SUPCON TECHNOLOGY CO., LTD.
No.309 Liuhe Road, BinJiang District, Hangzhou,Zhejiang, P.R.China
luwj@supcon.com