

Compatibility of Puppeteer + Cheerio Router Scraper

Summary

Your current scraping setup works only on routers with old-style, server-rendered HTML interfaces (like your PTCL DSL-226). It breaks on modern routers that use JavaScript SPAs, APIs, or authenticated endpoints.

Where It Works

Routers that render HTML directly:

- Older TP-Link
- Older D-Link
- Huawei HG series (old firmware)
- ZTE older models
- Netgear older models
- Any router using .asp, .htm, CGI, or simple HTML templates

Why it works:

- All data is inside the HTML.
- No dynamic SPA.
- No API tokens.
- No session-heavy requests.
- Cheerio can parse everything offline.

Where It Fails

Routers that use:

- React/Vue/Angular admin panels
- API-driven admin interfaces
- Token-based authenticated XHR requests
- Encrypted backend communication
- App-only configuration (Nest WiFi, new mesh systems)

Why it fails:

- No HTML to parse.
- Data loads through JS, not server-rendered.
- Requires cookies, tokens, or signatures.
- Cheerio doesn't execute JS.

What Determines Compatibility

Your agent only works if:

- The interface is HTML-first, not JS-first.
- Pages contain static markup.
- No dynamic API fetch is required.

Compatibility Table

Type of Router UI | Will It Work? | Reason

Static HTML pages | Yes | Cheerio can parse

ASP/JSP/PHP pages | Yes | Server-rendered HTML

CGI pages | Yes | Raw HTML exists

Modern JS SPA | Maybe | Requires DOM evaluation

API-only router | No | Needs token/API inspection

App-only router | No | No HTML interface exists

How To Make Your Agent Universal

1. Detect Page Type Automatically

- Load with Puppeteer
- Check for HTML vs JS SPA
- Check for XHR requests

- Decide scraping strategy

2. Add Router Profiles

Each router brand has predictable endpoints. Build adapters for:

- TP-Link
- D-Link
- ZTE
- Huawei
- Fiber ONTs

3. Use Network Interception

Puppeteer can capture:

- API responses
- AJAX calls
- Token flows

This bypasses the need for HTML scraping entirely.

Final Conclusion

Your current agent works only on old-school routers like PTCL DSL■226. Anything modern needs:

- DOM execution
- API reverse-engineering
- Network interception
- Or router-specific scraping profiles.