

Seeing the Unseen: Entropy-Guided Insights into Binary Files

Andrada-Livia Antoneac, Gheorghita Mutu, Dragoş-Teodor Gavriluţ

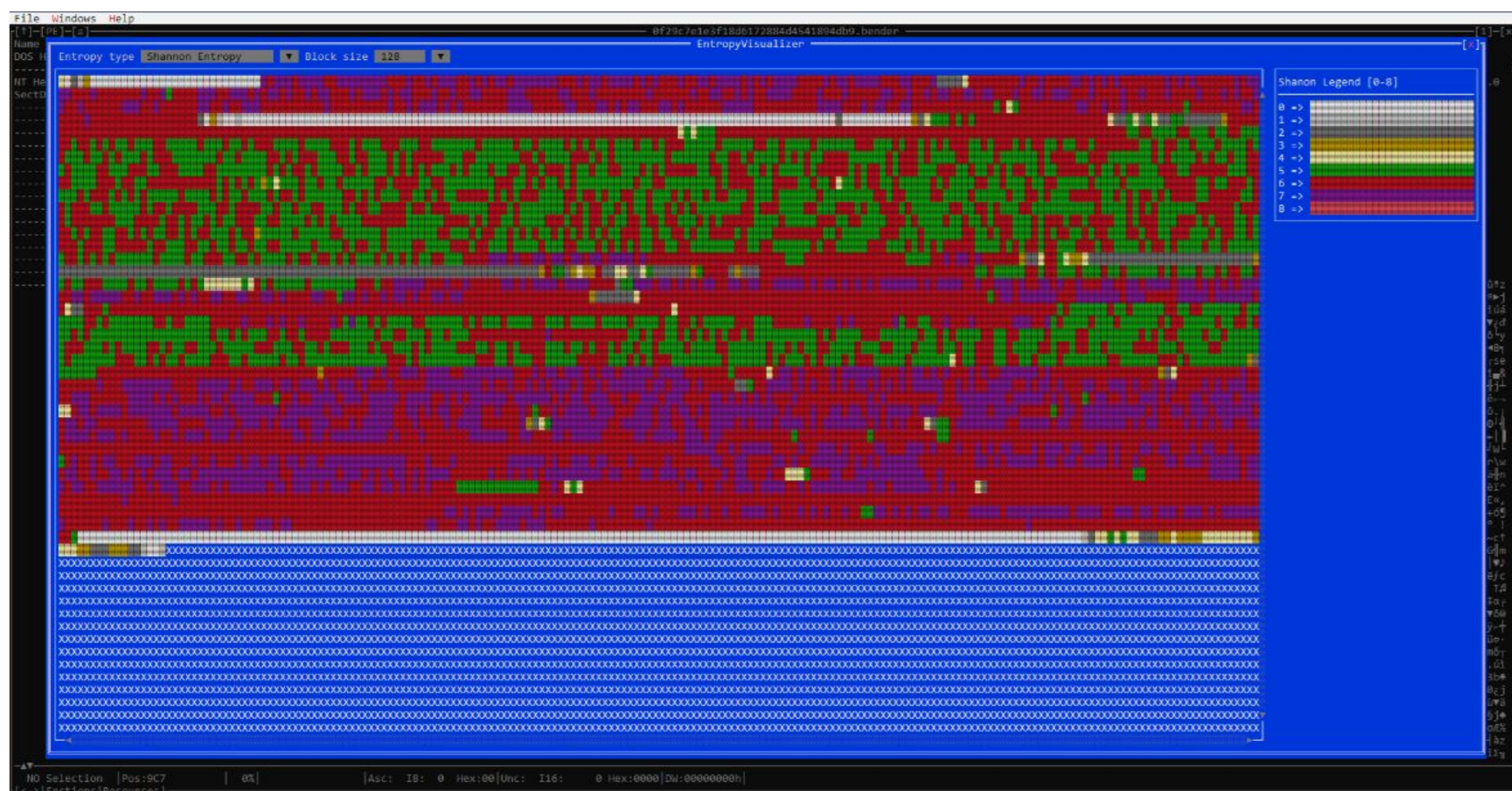


Why entropy?

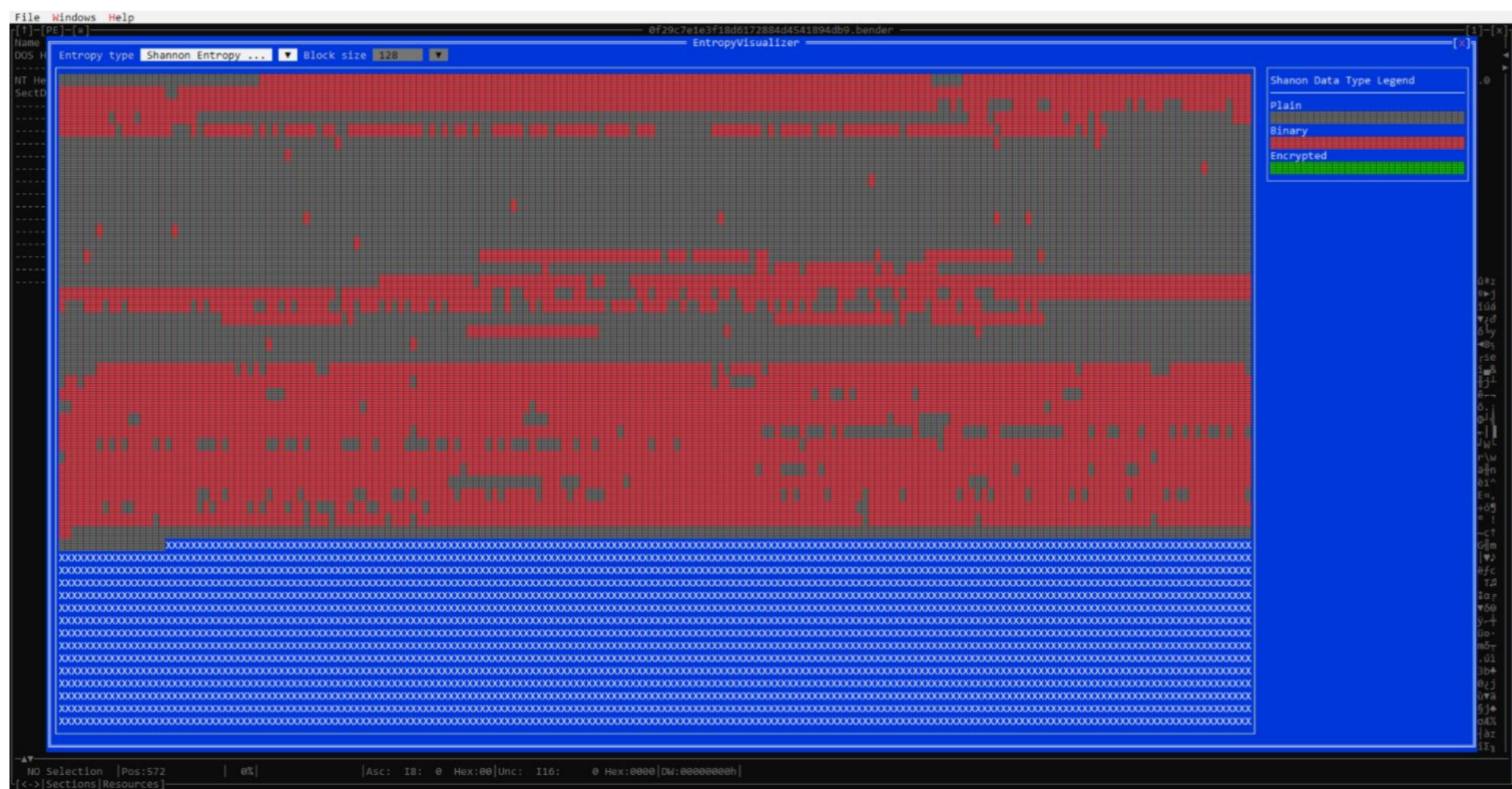
As software grows increasingly complex, so too do the vulnerabilities that malicious actors can exploit to compromise sensitive information or operations. Static analysis, a technique for examining code without execution, plays a crucial role in identifying flaws and risky patterns early in the development process, thereby enhancing code quality, reducing risk, and supporting compliance with security standards. Beyond traditional code review, analyzing the entropy of binary files offers valuable insights into potential obfuscation, encryption, or compression by measuring the randomness in byte value distributions.

Entropy Computation

To create a visual representation of a file, we integrated a new visualizer into GView, which computes the Shannon entropy for byte values within each block of bytes and associates it with a color presented in the legend.



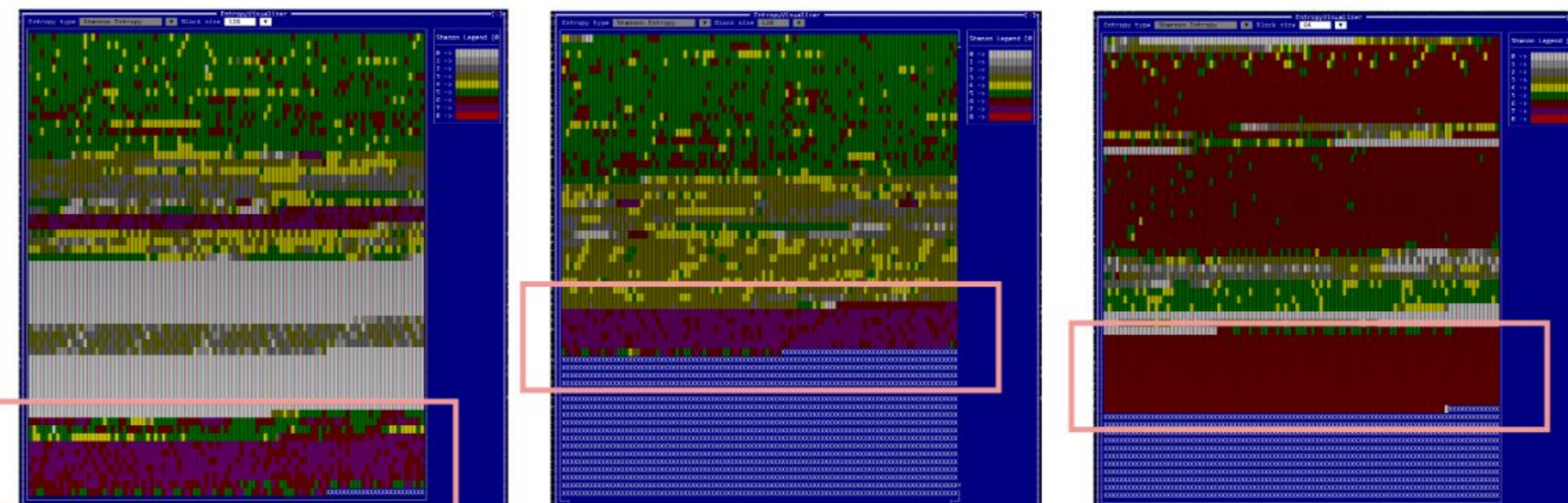
Granular entropy view



Entropy overview

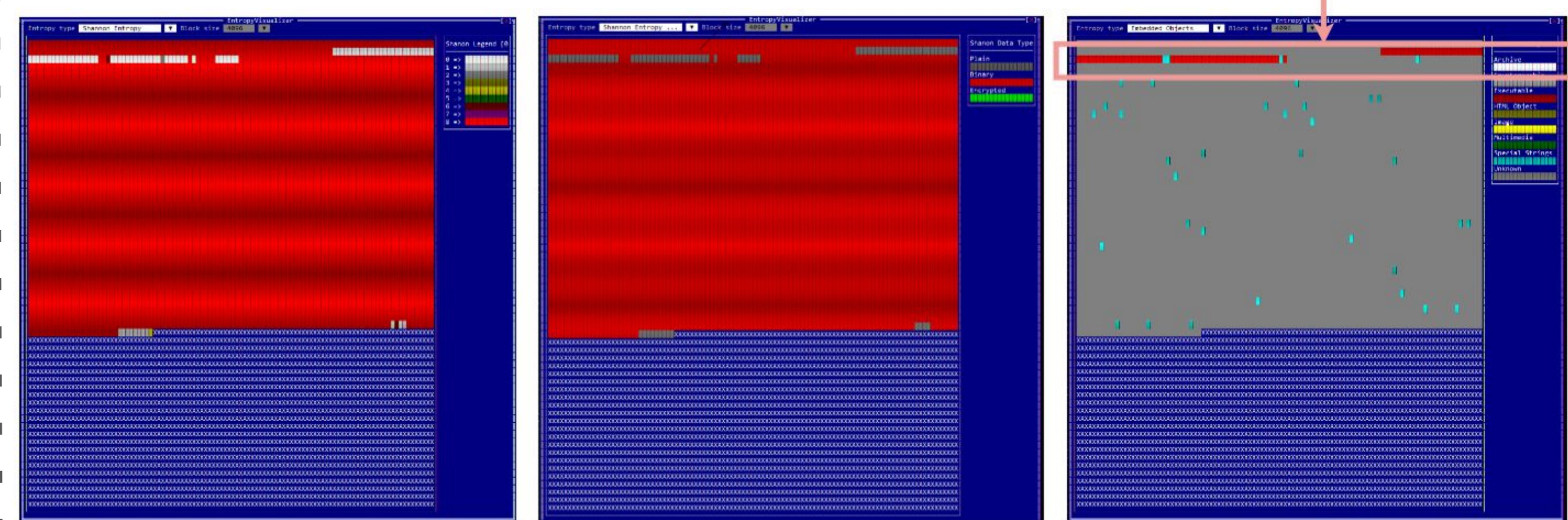
Separating plain text, binary data, and encrypted data was possible by defining three entropy value intervals, offering a high-level overview of the file.

Results



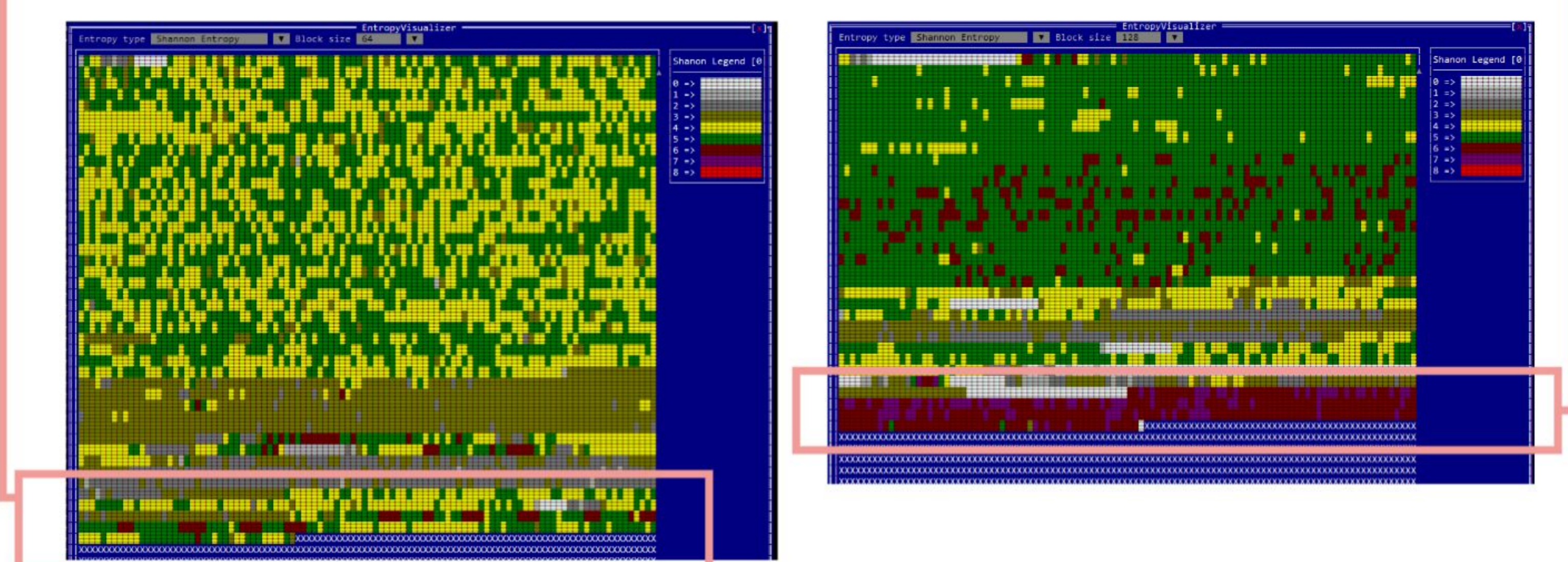
Files from the same malware family having similar entropy patterns

PNG file hiding a MZPE file



Executable hiding in an image file

Malicious payload in executable



Two files with malicious payloads - identified by the change in entropy values in uncommon areas

References

1. Raul Zaharia, Dragoş Gavriluţ, Gheorghita Mutu, Dorel Lucanu. 2024. GView: A Versatile Assistant for Security Researchers
2. Andrada-Livia Antoneac; Gheorghita Mutu; Dragoş-Teodor Gavriluţ. 2024. Entropy-Driven Visualization in GView: Unveiling the Unknown in Binary File Formats



12th ACM Celebration of Women in Computing: womENCourage™
Braşov, Romania
17-19 September, 2025
Theme: Computer Science: a Catalyst for Educational Change

