10B trace, egine slice se 30M traces kai ekei vrikame H2P, μετα καναμε το SimPoint. polla workloads. Αυτό εγινε για να δουμε πολλες διαφορετικες φασεις και πολλα διαφορετικα H2P.

Emeis exoume 2B trace to opoio einai apo to SimPoint, exoume mono ena workload. Δεν μπορουμε να κανουμε τοσο καλο screening, αλλα μπορουμε να εξαγουμε branches τα οποια εχουν πολλα misspredicts και μεγαλα occurrences

Όπως και στο paper το νευρωνικο είναι application specific. Πχ για datacenter. Ωστόσο δεν βασιζεται σε source code ή analysis, το analysis εχει γινει από πριν. Ισως να μπορεσουμε να κανουμε πιο general νευρωνικο και να ανιχνευει σε πιο πολλα benchmarks τα h2p. Στο paper δειχνουν ότι αν κανεις τραιν σε ένα benchmark με χ workload, εχεις καλο accuracy και σε y workload στο ιδιο benchmark. Δεν μπορουμε μεν να το τσεκαρουμε, αλλα βλεπουμε ότι με μικρο train, γινεται καλο accuracy σε πολύ μεγαλο validation για ένα benchmark.

Lstm γιατι εχουμε sequence, και ιστορια, ενώ τα CNN είναι καλα για spatial data.

Ισως δοκιμασαν ηδη το lstm και δεν τους δουλεψε τοσο καλα. Ισως επειδη το CNN είναι το πιο straightforward πραγμα που μπορεις να δοκιμασεις.

Αυτοι κανουν encode PC + TAKEN => 0-256 number σε πινακα 256χ200 1-hot. Και κανουν ετσι predict. Εγω προς το παρων δινω input μονο τα Boolean αλλα θα δοκιμασουμε και το άλλο.